

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 15:10:14 ; Search time 156 Seconds  
(without alignments)  
1027.897 Million cell updates/sec

Title: US-09-879-248-6

Perfect score: 2310

Sequence: 1 MSILILNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	AAW96260	Hypersens
2	2310	100.0	447	AAW71094	Erwinia a
3	2310	100.0	447	AAW84855	A hyperse
4	2310	100.0	447	AAO22548	Hypersens
5	2310	100.0	447	AAE18296	Erwinia a
6	2310	100.0	447	AAE16448	E. amylov
7	2310	100.0	447	ABO92226	Erwinia a
8	559	24.2	424	AAW97851	Hypersens
9	559	24.2	424	AAW71098	Pseudomon
10	559	24.2	424	AAW83014	Pseudomon
11	559	24.2	424	AAW84859	A hyperse
12	559	24.2	424	AAO22550	Hypersens
13	559	24.2	424	AAE18298	Pseudomon
14	559	24.2	424	AAE16452	P. syring
15	359.5	15.6	197	ABO92228	P. syring
16	359.5	15.6	197	AAW83014	Bacillus
17	341	14.8	221	AAW77412	Bacillus
18	341	14.8	221	AAW28446	Bacillus
19	199.5	8.6	1306	ADK48880	Bacillus
20	193	8.4	694	ABU36481	Protein e
21	191.5	8.3	584	ABU36802	Protein e
22	191	8.3	2280	ABU22680	Protein e
23	188	8.1	639	ABU36685	Protein e
24	186.5	8.1	591	ABU36439	Protein e
25	185.5	8.0	1381	ABU36971	Protein e

26	185	8.0	588	7	ADM40790	Adm40790 Mycobacte
27	183	7.9	588	6	ABU36945	Protein e
28	182	7.9	606	6	ABU36540	Protein e
29	181.5	7.9	484	6	ABU15873	Mycobacte
30	181.5	7.9	484	6	ABU15873	Mycobacte
31	181.5	7.9	484	6	ABU15873	Mycobacte
32	181.5	7.9	484	6	ABU15873	Mycobacte
33	180.5	7.8	1079	3	ABO23518	Haemophil
34	179	7.7	730	3	AAO23519	Mycobacte
35	179	7.7	730	3	AAO23519	Mycobacte
36	176.5	7.6	562	6	ABU36590	Protein e
37	176.5	7.6	2090	4	ABU36590	Protein e
38	174	7.5	615	6	ABU36862	Protein e
39	174	7.5	2042	2	AAW56319	Haemophil
40	174	7.5	2586	4	ABU36878	Drosophil
41	172.5	7.5	584	6	ABU36979	Protein e
42	171.5	7.4	667	6	ABU36570	Protein e
43	171	7.4	505	6	ABU34417	Protein e
44	169.5	7.3	914	6	ABU36682	Protein e
45	169	7.3	525	6	ABU36839	Protein e

## ALIGNMENTS

RESULT 1  
AAW96260  
ID AAW96260 standard; protein; 447 AA.  
XX  
AC AAW96260;  
XX  
DT 14-JUN-1999 (first entry)  
XX  
DE Hypersensitive response eliciting protein (HrpW).  
XX  
KW Hypersensitive response elicitor protein; hypersensitive response; hrpW;  
KW pathogen; infection; crop protection; disease resistance;  
KW pest resistance; transgenic plant; colouration; maturation.  
XX  
OS Erwinia amylovora.  
XX  
PN WO9907208-A1.  
XX  
PD 18-FEB-1999.  
XX  
PF 27-JUL-1998; 98WO-US015547.  
XX  
PR 06-AUG-1997; 97US-0055108P.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI Kim JF, Beer SV;  
XX  
DR WPI; 1999-167126/14.  
XX  
DR N-PSDB; AAX09007.  
XX  
PT New Erwinia amylovora hypersensitive response eliciting gene and protein  
XX  
PT - useful for providing transgenic plants and seeds with enhanced growth,  
XX  
XX and insect and disease resistance.  
XX  
PS Claim 1; Page 50-51; 54pp; English.  
XX  
CC The hypersensitive response eliciting protein (hrp) or polypeptide is  
XX  
CC produced as part of an active defense by plants against incompatible  
XX  
CC pathogen infections. The hypersensitive response is a rapid localised  
XX  
CC necrosis. The hrp protein and gene when used in nucleotide constructs are  
XX  
CC useful for providing disease resistance to plants, insect control to  
XX  
CC plants, and enhancing plant growth (enhancing fruit size and earlier  
XX  
CC colouration and maturation), by direct application of the protein to  
XX  
XX plants, or by producing transgenic plants or seeds using the hrp gene  
XX  
SQ Sequence 447 AA;

Query Match 100.0%; Score 2310; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHHNANSALGQOPIIDROTIEQMAQLLAELLKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHHNANSALGQOPIIDROTIEQMAQLLAELLKSLLS 60

QY 61 POSGNAATGAGNDQTTGTVGNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGTVGNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGGLDQAITPDG 120

QY 121 QGGQIGDNPILLKAMLLIARMDGQDQFGQPGTGNNSASSTSSGSGSPFNDLSGGKA 180  
DB 121 QGGQIGDNPILLKAMLLIARMDGQDQFGQPGTGNNSASSTSSGSGSPFNDLSGGKA 180

QY 181 PSGNSPSGNYSPVSTFPPSTPTSPDLPSSPTKAAGGTPVTDHPDPVGSAGIGAG 240  
DB 181 PSGNSPSGNYSPVSTFPPSTPTSPDLPSSPTKAAGGTPVTDHPDPVGSAGIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDGAS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDGAS 300

QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360  
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360

QY 361 KILQNLADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKSDSEGLN 420  
DB 361 KILQNLADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKSDSEGLN 420

QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 3  
AAY84855  
ID AAY84855 standard; protein; 447 AA.  
XX AAY84855;  
XX  
XX 08-AUG-2000 (first entry)  
XX  
XX A hypersensitive response elicitor protein.  
XX  
XX Hypersensitive response; insect control; disease resistance;  
XX KW Hypersensitive response elicitor; plant growth; vegetable; crop;  
XX KW ornamental plant.  
XX  
XX Erwinia amylovora.  
XX  
XX WO200020452-A2.  
XX  
XX 13-APR-2000.  
XX  
XX 05-OCT-1999; 99WO-US023181.  
XX  
XX 05-OCT-1998; 98US-0103050P.  
XX PR (EDEN-) EDEN BIOSCIENCE CORP.  
XX PA  
XX Wei Z, Fan H, Niggenmeyer JL;  
XX WPI; 2000-303745/26.  
XX DR

Query Match 100.0%; Score 2310; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHHNANSALGQOPIIDROTIEQMAQLLAELLKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHHNANSALGQOPIIDROTIEQMAQLLAELLKSLLS 60

QY 61 POSGNAATGAGNDQTTGTVGNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGTVGNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGGLDQAITPDG 120

QY 121 QGGQIGDNPILLKAMLLIARMDGQDQFGQPGTGNNSASSTSSGSGSPFNDLSGGKA 180  
DB 121 QGGQIGDNPILLKAMLLIARMDGQDQFGQPGTGNNSASSTSSGSGSPFNDLSGGKA 180

QY 181 PSGNSPSGNYSPVSTFPPSTPTSPDLPSSPTKAAGGTPVTDHPDPVGSAGIGAG 240  
DB 181 PSGNSPSGNYSPVSTFPPSTPTSPDLPSSPTKAAGGTPVTDHPDPVGSAGIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDGAS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDGAS 300

QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360  
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360

QY 361 KILQNLADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKSDSEGLN 420  
DB 361 KILQNLADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKSDSEGLN 420

QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 2  
AAY71094  
ID AAY71094 standard; protein; 447 AA.  
XX AAY71094;  
XX  
XX 08-SEP-2000 (first entry)  
XX  
XX Erwinia amylovora hypersensitive response elicitor #2.  
XX  
XX Hypersensitive response elicitor; environmental stress resistance; plant.  
XX  
XX Erwinia amylovora.  
XX  
XX WO200028055-A2.  
XX  
XX 18-MAY-2000.  
XX  
XX 04-NOV-1999; 99WO-US026039.  
XX  
XX 05-NOV-1998; 98US-0107243P.  
XX PR (EDEN-) EDEN BIOSCIENCE CORP.  
XX PA  
XX Wei Z, Schading RL;  
XX  
XX WPI; 2000-376566/32.  
XX DR N-PSDB; AAD00669.  
XX  
XX Application of a hypersensitive response elicitor protein to plants to  
XX PT impart stress resistance.  
XX  
XX Disclosure; Page 10-12; 84pp; English.  
XX  
XX The patent discloses a method to impart stress resistance to plants by  
XX CC applying a hypersensitive response elicitor in a non-infectious form to a

DR N-PSDB; AAL4939.  
XX Hypersensitive response elicitor polypeptides useful for imparting  
PT enhanced growth, disease resistance and insect resistance to plants,  
PT especially vegetables and ornamental flowers.  
XX  
PS Disclosure; Page 12-13; 100pp; English.  
XX  
XX The present sequence represents a hypersensitive response elicitor  
CC polypeptide. The specification describes hypersensitive response elicitor  
CC polypeptide fragments, which do not elicit a hypersensitive response.  
CC Instead, the proteins impart disease resistance to plants, enhance plant  
CC growth, and/or control insects. The polypeptide fragments may be used to  
CC these properties to plants. The plants which may be treated in this way  
CC include vegetables, crops and ornamental plants such as alfalfa, rice,  
CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet  
CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,  
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,  
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,  
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,  
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis  
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,  
CC carnation or zinnia  
XX  
SQ Sequence 447 AA;  
Query Match 100.0%; Score 2310; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNNTSSPGLFQSGGNGGLGHNANSALGQOPIDRTIQEOMALLAELLKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGGNGGLGHNANSALGQOPIDRTIQEOMALLAELLKSLLS 60  
QY 61 POSGNAATGAGNDQTTGVNAGGLNGRGTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGVNAGGLNGRGTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
QY 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGOPGTGNNSSASGTSSSGSGSPFNDLSGGKA 180  
DB 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGOPGTGNNSSASGTSSSGSGSPFNDLSGGKA 180  
QY 181 PSGNSPBGNYSPVSTFSPSTPTSPPLDFFSSPTKAAAGGSTPVTDPDVGSAIGAG 240  
DB 181 PSGNSPBGNYSPVSTFSPSTPTSPPLDFFSSPTKAAAGGSTPVTDPDVGSAIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITTVKAGQVFDGKGTFTAGSELGQGGQSENOKPLFILEDGS 300  
DB 241 NSVAFTSAGANQTVLHDTITTVKAGQVFDGKGTFTAGSELGQGGQSENOKPLFILEDGS 300  
QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
QY 361 KILQNLNADTNLSVDNVKADFGFTVNTGQQQGNWDLNLSHISAEDKGFVFKSDSEGLN 420  
DB 361 KILQNLNADTNLSVDNVKADFGFTVNTGQQQGNWDLNLSHISAEDKGFVFKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
RESULT 4  
ID AAO22548 standard; protein; 447 AA.  
XX AAO22548;  
AC AAO22548;  
XX  
DT 28-OCT-2002 (first entry)  
XX  
DE Hypersensitive response elicitor protein #2 from Erwinia amylovora.  
XX

KW Desiccation inhibitor; longevity enhancer; desiccation inhibitor;  
KW cutting; ornamental plant; hypersensitive response elicitor protein;  
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;  
KW flower bloom; flower.  
XX  
OS Erwinia amylovora.  
XX  
PN WO200237960-A2.  
XX  
PD 16-MAY-2002.  
XX  
XX 06-NOV-2001; 2001WO-US043715.  
XX  
PR 13-NOV-2000; 2000US-0248169P.  
XX  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
XX  
PI Wei Z, Leon E, Oviedo A;  
XX  
DR WPI; 2002-575194/61.  
DR N-PSDB; AAL41133.  
XX  
PT Inhibiting desiccation of cuttings from ornamental plants, by treating  
PT ornamental plants with hypersensitive response elicitor protein, or  
PT expressing heterologous hypersensitive response elicitor protein in  
PT plants.  
XX  
PS Disclosure; Page 12-13; 69pp; English.  
XX  
XX The invention relates to a method for inhibiting desiccation of cuttings  
CC from ornamental plants. The method involves treating the cuttings with a  
CC hypersensitive response elicitor protein or polypeptide, or providing a  
CC transgenic ornamental plant or plant seed transformed with a DNA molecule  
CC encoding the hypersensitive response elicitor polypeptide, and growing  
CC the ornamental plant or transgenic ornamental plant produced from the  
CC transgenic ornamental plant seed. The hypersensitive response elicitor  
CC protein or polypeptide is useful for inhibiting desiccation of cuttings  
CC from ornamental plants, for harvesting cuttings from ornamental plants,  
CC for promoting early flowering of ornamental plants, and enhancing the  
CC longevity of flower blooms on ornamental plant cuttings. This sequence  
CC represents a hypersensitive response elicitor protein of the invention  
XX  
SQ Sequence 447 AA;  
Query Match 100.0%; Score 2310; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSILTLNNNTSSPGLFQSGGNGGLGHNANSALGQOPIDRTIQEOMALLAELLKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGGNGGLGHNANSALGQOPIDRTIQEOMALLAELLKSLLS 60  
QY 61 POSGNAATGAGNDQTTGVNAGGLNGRGTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGVNAGGLNGRGTAGTTPOSQSNMLSEMGNGLDQAITPDG 120  
QY 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGOPGTGNNSSASGTSSSGSGSPFNDLSGGKA 180  
DB 121 QGGGQIGDNPLLKAMLKLIARMDGQSDQFGOPGTGNNSSASGTSSSGSGSPFNDLSGGKA 180  
QY 181 PSGNSPBGNYSPVSTFSPSTPTSPPLDFFSSPTKAAAGGSTPVTDPDVGSAIGAG 240  
DB 181 PSGNSPBGNYSPVSTFSPSTPTSPPLDFFSSPTKAAAGGSTPVTDPDVGSAIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITTVKAGQVFDGKGTFTAGSELGQGGQSENOKPLFILEDGS 300  
DB 241 NSVAFTSAGANQTVLHDTITTVKAGQVFDGKGTFTAGSELGQGGQSENOKPLFILEDGS 300  
QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
QY 361 KILQNLNADTNLSVDNVKADFGFTVNTGQQQGNWDLNLSHISAEDKGFVFKSDSEGLN 420  
DB 361 KILQNLNADTNLSVDNVKADFGFTVNTGQQQGNWDLNLSHISAEDKGFVFKSDSEGLN 420

Db	361	KIQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN	420
Qy	421	VNTSDISLGDVENHYKVPMSANLKVAE	447
Db	421	VNTSDISLGDVENHYKVPMSANLKVAE	447
RESULT 5			
AAE18296			
ID	AAE18296 standard; protein; 447 AA.		
XX	AC	AAE18296;	
XX	DT	07-MAY-2002 (first entry)	
XX	DE	Erwinia amylovora hypersensitive response elicitor (HRE) #2.	
XX	KW	Hypersensitive response elicitor; HRE; transgenic plant; plant growth; stress tolerance; disease tolerance; modified flower colour; insect resistance; herbicide resistance; male sterility.	
XX	KW	insect resistance; herbicide resistance; male sterility.	
XX	OS	Erwinia amylovora.	
XX	PN	WO200195724-A2.	
XX	PD	20-DEC-2001.	
XX	PF	13-JUN-2001; 2001WO-US018955.	
XX	PR	15-JUN-2000; 2000US-0211585P.	
XX	PA	(EDEN-) EDEN BIOSCIENCE CORP.	
XX	PI	Wei Z, Derocher J;	
XX	DR	WPI; 2002-130707/17.	
XX	DR	N-PSDB; AAD29125.	
XX	CC	Improving effectiveness of transgenic plants by topical application of a hypersensitive response elicitor protein to the transgenic plant or by incorporating into the plant a transgene encoding the protein.	
XX	PS	Disclosure; Page 13-14; 86pp; English.	
XX	CC	The invention relates to methods of improving the effectiveness of transgenic plants which involves either topical application of a hypersensitive response elicitor (HRE) protein to the transgenic plant or incorporating into the transgenic plant a transgene encoding HRE. HRE sequence is used for improving the effectiveness of transgenic plants by maximising the benefit of transgenic traits associated with a deleterious effect on growth, stress tolerance, disease or insect resistance, enhanced growth, herbicide resistance, male sterility, modified flower colour and biochemically modified plant product in the transgenic plants or overcoming the deleterious effects. The present sequence is Erwinia amylovora HRE protein	
XX	SQ	Sequence 447 AA;	
Query Match			
Best Local Similarity 100.0%; Score 2310; DB 5; Length 447;			
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MSILTLNNNTSSPGLFQSGGDNGLGHNANALGOQPIDROTIEQMAQLLAELKSLLS	60
Db	1	MSILTLNNNTSSPGLFQSGGDNGLGHNANALGOQPIDROTIEQMAQLLAELKSLLS	60
Qy	61	PQSGNAATGAGGNDQTTGVNAGGLNGRGTAGTTTPQSDSQNMLSEMGNGLDQAITPDG	120
Db	61	PQSGNAATGAGGNDQTTGVNAGGLNGRGTAGTTTPQSDSQNMLSEMGNGLDQAITPDG	120
Qy	121	QGGQIGDNPLLKAMLKLIARMDGQSDQFGQPGTGNNNSASGTSSTSSGSPNDLSGGKA	180

PS Disclosure; Page 13-14; 99pp; English.

XX The patent discloses hypersensitive response elicitor proteins and nucleotides encoding such proteins. Hypersensitive response elicitor proteins comprise an isolated pair or more of spaced apart domains, each comprising an acidic portion linked to an alpha-helix and capable of eliciting a hypersensitive response in plants. Sequences of the invention are used to impart disease resistance to plants, to enhance plant growth, to control insects and/or to impart stress resistance to plants which includes resistance to environmental stresses such as climate, air pollution, chemical and nutritional stress. The method of imparting disease resistance has the potential for treating previously untreatable diseases, treating diseases systemically and avoiding the use of infectious agents or environmentally harmful materials. Hyper- sensitive response elicitor sequences are used to enhance plant growth which encompasses greater yield, increased in quantity of seeds produced, percentage of seeds germinated, plant size and biomass, bigger fruits, earlier fruit coloration and plant maturation. They are also used for insect control which encompasses preventing direct insect damage to plant by feeding injury, interfering with insect larval feeding on the plants, preventing insects from colonizing host plants and releasing phytotoxins. Sequences of the invention also prevent subsequent disease damage to plants resulting from insect infection. The present sequence is Erwinia amylovora hypersensitive response elicitor protein

XX Sequence 447 AA;

Query Match 100.0%; Score 2310; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIRDTQTIQMAQLLAKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIRDTQTIQMAQLLAKSLLS 60

QY 61 POSGNAATGAGNDQTTGVNAGLNGRGTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGVNAGLNGRGTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120

QY 121 QGGGQIGDNPLLKAMLKLIARMDGSDQFGPGTGNNSSAGTSSSGGSPFNDLSGGKA 180  
DB 121 QGGGQIGDNPLLKAMLKLIARMDGSDQFGPGTGNNSSAGTSSSGGSPFNDLSGGKA 180

QY 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAGGSTPVTDHPDVGSAIGAG 240  
DB 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAGGSTPVTDHPDVGSAIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQGGQSENKPLFILEDGS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQGGQSENKPLFILEDGS 300

QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVITNSSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVITNSSFEHSD 360

QY 361 KILQLNADTNLSVDNVKADFGFTVTNGGQQGNWDLNLSHISAEDKFSFVKSDSEGLN 420  
DB 361 KILQLNADTNLSVDNVKADFGFTVTNGGQQGNWDLNLSHISAEDKFSFVKSDSEGLN 420

QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 7  
ABB09226  
ID ABB09226 standard; protein; 447 AA.  
XX  
AC ABB09226;  
XX  
DT 08-JUL-2002 (first entry)  
XX  
DE Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.

XX Hypersensitive response elicitor; fruit; vegetable; plant; desiccation; postharvest disease.  
XX Erwinia amylovora.  
XX WO200180639-A2.  
XX 01-NOV-2001.  
XX 17-APR-2001; 2001WO-US012468.  
XX 19-APR-2000; 2000US-0198359P.  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
XX Wei Z, Qiu D, Remick D;  
XX WPI: 2002-041357/05.  
XX N-PSDB; ABUS1711.  
XX Inhibiting post harvest disease (caused by Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a fruits or vegetables, using hypersensitive response elicitor proteins or nucleic acids.  
XX Disclosure; Page 11-12; 72pp; English.  
XX The present invention describes methods for inhibiting post harvest disease or desiccation and enhancing the longevity in a fruits or vegetables, using hypersensitive response elicitor proteins or polypeptides or nucleic acids (I) derived from pathogens (e.g. Erwinia amylovora, E. Stewartii, E. chrysanthemi, E. carotovora, Xanthomonas, Pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter). (I) has bactericidal activity, and can be used in gene therapy. The method can be used for inhibiting post harvest disease (caused by Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a fruits or vegetables. The method enables growers, warehouse packers, shippers and suppliers to process, handle and store fruit and vegetables with reduced losses caused by post harvest disease and desiccation, therefore reducing costs to the consumer and improving quality. The present sequence represents a hypersensitive response elicitor protein given in the exemplification of the present invention  
XX Sequence 447 AA;  
Query Match 100.0%; Score 2310; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIRDTQTIQMAQLLAKSLLS 60  
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIRDTQTIQMAQLLAKSLLS 60

QY 61 POSGNAATGAGNDQTTGVNAGLNGRGTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGVNAGLNGRGTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120

QY 121 QGGGQIGDNPLLKAMLKLIARMDGSDQFGPGTGNNSSAGTSSSGGSPFNDLSGGKA 180  
DB 121 QGGGQIGDNPLLKAMLKLIARMDGSDQFGPGTGNNSSAGTSSSGGSPFNDLSGGKA 180

QY 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAGGSTPVTDHPDVGSAIGAG 240  
DB 181 PSNGSPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAGGSTPVTDHPDVGSAIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQGGQSENKPLFILEDGS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGQGGQSENKPLFILEDGS 300

QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVITNSSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVITNSSFEHSD 360

Db 301 LKNVTWGGDAGDIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVITNSSFEHSD 360  
QY 361 KILQLNADTNLSVDNKAQDFGFTVNTGSGQGNWDLNLSHISAEDGKSFVKSDSEGLN 420  
Db 361 KILQLNADTNLSVDNKAQDFGFTVNTGSGQGNWDLNLSHISAEDGKSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 8

AAW97851  
ID AAW97851 standard; protein; 424 AA.  
AC AAW97851;  
DT 07-JUN-1999 (first entry)  
DE Hypersensitive response elicitor HrpW.  
XX  
KW Hypersensitive response elicitor; dspE gene; HrpW; transgenic plant;  
KW disease resistance; insect resistance.  
XX  
OS Pseudomonas syringae pv. tomato.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..186  
FT Region 119..186  
FT /note= "hypersensitive response elicitor-like domain"  
FT /note= "region of 6 imperfect glycine-rich repeats"  
FT Domain 187..424  
FT /note= "C-terminal domain"  
XX  
PN WO9907207-A1.  
XX  
PD 18-FEB-1999.  
XX  
PF 24-JUL-1998; 98WO-US015501.  
XX  
PR 06-AUG-1997; 97US-0055107P.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI Collmer A, Charkowski A, Alfano JR;  
XX  
DR WPI; 1999-167125/14.  
DR N-PSDB; AAX24347.  
XX  
PT New hypersensitive response eliciting (dspE) gene and protein - useful  
PT for providing transgenic plants and seeds with enhanced growth, and  
PT insect and disease resistance.  
XX  
PS Claim 18; Page 38-39; 56pp; English.  
XX  
CC This the amino acid sequence of hypersensitive response element (HRE)  
CC HrpW of Pseudomonas syringae pv. tomato DC3000, as deduced from the  
CC nucleotide sequence of an isolated DNA molecule (see AAX24347). HrpW  
CC includes an HRE-like domain that is rich in Gln, Ser and Gly and which  
CC includes 6 imperfect glycine-rich repeats with many acidic and polar  
CC residues that align with similar repeats in the HrpZ proteins of P.  
CC syringae pv. syringae and P. syringae pv. tomato. The amino acid sequence  
CC of this region suggests alternating beta-sheets and turns that may form a  
CC barrel structure. The C-terminal domain of HrpW is similar to several  
CC fungal and bacterial Pel proteins. The HrpW protein or isolated DNA  
CC molecule can be used to impart disease resistance to plants, to enhance  
CC plant growth and/or to control insects on plants. This is achieved by  
CC applying the HRE protein in a non-infectious form to plants or plant  
CC seeds. Alternatively, transgenic plants or plant seeds transformed with  
CC DNA encoding the HRE can be provided  
XX  
SQ Sequence 424 AA;

Query Match 24.2%; Score 559; DB 2; Length 424;  
Best Local Similarity 36.1%; Pred. No. 2.7e-30;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
QY 37 QPIDQTTEQMAQLLAELLKSL---LSPQSGNAATGAGGNDOTTGVGNAGGLNKRKGTAG 93  
Db 72 KENDSQS--NTAKLISALIMSLQLMLTNSNKKQDTNORQPDQSQAPFQNNGLG----- 122  
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGOGGQIGDNPLLKAMLLKLIARMMDGSDQFGQP 153  
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151  
QY 154 GTGNNSASGTSSTSSGCGSPFNLSGKAPSGNSPSPVSTFSPPTSPSTPSLDFPS 213  
Db 152 GGDTPATGGGGGGGGTPTATGGG---SGGTPATGGGGGGVTQITPQL-----A 200  
QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFSTAGANQTVLHDTITVKAGVFGKGGT 273  
Db 201 NPNTSG-----TGSVSDTAGS---TEQAGKINVVKOTIKVGAGEVFDHGAT 245  
QY 274 FTAGSELGDDGSGSENOKPLFILEDGASLKNVTMGDDGADGIHLYG---DAKIDNLHVTN 329  
Db 246 FTADKSMGNGDQGENQKPFELAEATLKNVLNKGNEVDGIHVAKAKNAQEVTTIDNVHAQN 305  
QY 330 VGEDAITVKPNSAGKSHVEITNSSFEHASDKILQLNADTNLSVDNKAQDFGFTVNTG 389  
Db 306 VGEDLITVKEGGAATVNLINKNSAKGADKVVQLNANTHLKIDNFKADDFGTWVRTNG 365  
QY 330 GQO-GNWDNLNLSHISAEDGKSFVKSDSEGLNVTSDISLGDVENHY 435  
Db 366 GKQFDDMSIELNGTEANHGKFPALVKSDSDLLKLATNTAMTVDVKHAY 412

RESULT 9

AAAY71098  
ID AAY71098 standard; protein; 424 AA.  
XX  
AC AAY71098;  
DT 08-SEP-2000 (first entry)  
DE Pseudomonas syringae dspE gene encoded hypersensitive response elicitor.  
KW Hypersensitive response elicitor; environmental stress resistance; plant;  
KW pathogen; dspE gene.  
XX  
OS Pseudomonas syringae.  
XX  
PN WO200028055-A2.  
XX  
PD 18-MAY-2000.  
XX  
PF 04-NOV-1999; 99WO-US026039.  
PR 05-NOV-1998; 98US-0107243P.  
PA (EDEN-) EDEN BIOSCIENCE CORP.  
PI Wei Z, Schading RL;  
XX  
DR WPI; 2000-376566/32.  
DR N-PSDB; AAD00673.  
XX  
PT Application of a hypersensitive response elicitor protein to plants to  
PT impart stress resistance.  
XX  
PS Disclosure; Page 25-26; 84pp; English.  
CC The patent discloses a method to impart stress resistance to plants by  
CC applying a hypersensitive response elicitor in a non-infectious form to a  
CC plant or seed. The present sequence is a hypersensitive response elicitor  
CC protein encoded by dspE gene from Pseudomonas syringae. The protein  
CC elicits a plant pathogen's hypersensitive response and is used to impart

```
CC stress resistance to plants
XX Sequence 424 AA;
SQ
  Query Match      24.2%; Score 559; DB 3; Length 424;
  Best Local Similarity 36.1%; Pred. No. 2.7e-30;
  Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL----LSPOSQNAATGAGGNDQTTGVGNAGGLNKRKGTAG 93
Db 72 KENDSQS--NIAKLISALIMSLLOMLTNSNKKQDTNQEQPDSQAPFQNNGLG-----122
QY 94 TTPQSDSQNMLEMNGNGLDQAITPDGQGGQIGDNPLLKMLKLIARMWDSQDFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNSASGTSSTSSGSPFNLDLSGKAPSGNPSGNYSVPSTFSPSTPTSPPLDFFS 213
Db 152 GDTPTATGGGGGGGTPTATGGG---SGGTPATGGGEGGVTPQITPQL-----A 200
QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFTSAGANTVLHDTITVKAGQVFDGKGT 273
Db 201 NPNRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVFDGHGAT 245
QY 274 FTAGSELGDSQSENOKPLFILEDGASLKNVTMGDDGADGHIHYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENOKPMFELAEAGATLKNVNLGENEVDGIHVAKNAQAEVTDINVAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSFEHASDKILOLNADTNLSVDNVKAKDFTGFTVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLINIKNSAKGADDKVQVQLNANTHLKIDNFKADDFGTWRTNG 365
QY 390 GQO-GNWDNLNLHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFKALVKSDSDDLKATGNIAMTDVKHAY 412

RESULT 10
AAV84859
ID AAV84859 standard; protein; 424 AA.
XX
AC AAV84859;
XX
DT 08-AUG-2000 (first entry)
DE A hypersensitive response elicitor protein.
XX
KW Hypersensitive response; insect control; disease resistance;
KW hypersensitive response elicitor; plant growth; vegetable; crop;
KW ornamental plant.
XX
OS Pseudomonas syringae.
XX
PN WO200020452-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US023181.
XX
PR 05-OCT-1998; 98US-0103050P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Fan H, Niggemeyer JL;
XX
XX WPI; 2000-303745/26.
XX
DR N-PSDB; AAA14943.
XX
PT Hypersensitive response elicitor polypeptides useful for imparting
PT enhanced growth, disease resistance and insect resistance to plants,
PT especially vegetables and ornamental flowers.
XX
PS Disclosure; Page 26-28; 100pp; English.
```

```
XX The present sequence represents a hypersensitive response elicitor
CC polypeptide. The specification describes hypersensitive response elicitor
CC polypeptide fragments, which do not elicit a hypersensitive response.
CC Instead, the proteins impart disease resistance to plants, enhance plant
CC growth, and/or control insects. The polypeptide fragments may be used to
CC these properties to plants. The plants which may be treated in this way
CC include vegetables, crops and ornamental plants such as alfalfa, rice,
CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet
CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,
CC carnation or zinnia
XX
SQ Sequence 424 AA;
  Query Match      24.2%; Score 559; DB 3; Length 424;
  Best Local Similarity 36.1%; Pred. No. 2.7e-30;
  Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL----LSPOSQNAATGAGGNDQTTGVGNAGGLNKRKGTAG 93
Db 72 KENDSQS--NIAKLISALIMSLLOMLTNSNKKQDTNQEQPDSQAPFQNNGLG-----122
QY 94 TTPQSDSQNMLEMNGNGLDQAITPDGQGGQIGDNPLLKMLKLIARMWDSQDFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNSASGTSSTSSGSPFNLDLSGKAPSGNPSGNYSVPSTFSPSTPTSPPLDFFS 213
Db 152 GDTPTATGGGGGGGTPTATGGG---SGGTPATGGGEGGVTPQITPQL-----A 200
QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFTSAGANTVLHDTITVKAGQVFDGKGT 273
Db 201 NPNRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVFDGHGAT 245
QY 274 FTAGSELGDSQSENOKPLFILEDGASLKNVTMGDDGADGHIHYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENOKPMFELAEAGATLKNVNLGENEVDGIHVAKNAQAEVTDINVAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSFEHASDKILOLNADTNLSVDNVKAKDFTGFTVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLINIKNSAKGADDKVQVQLNANTHLKIDNFKADDFGTWRTNG 365
QY 390 GQO-GNWDNLNLHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFKALVKSDSDDLKATGNIAMTDVKHAY 412

RESULT 11
AAO22550
ID AAO22550 standard; protein; 424 AA.
XX
AC AAO22550;
XX
DT 28-OCT-2002 (first entry)
DE Hypersensitive response elicitor protein #2 from Pseudomonas syringae.
XX
KW Desiccation inhibitor; longevity enhancer; desiccation inhibitor;
KW cutting; ornamental plant; hypersensitive response elicitor protein;
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;
KW flower bloom; flower.
XX
OS Pseudomonas syringae.
XX
PN WO200237960-A2.
XX
PD 16-MAY-2002.
XX
```





366 GKQFDDMSIELNGIEANHGKFAVLRSDSDLLKATGNIAMTDVKHAY 412

Db

RESULT 13  
AAE16452  
ID AAE16452 standard; protein; 424 AA.

AC AAE16452;

DT 09-APR-2002 (first entry)

DE P. syringae hypersensitive response elicitor protein, dspE.

KW Hypersensitive response elicitor protein; plant growth; fruit coloration;  
KW disease resistance; stress resistance; phytotoxin; insect infection;  
KW plant maturation; dspE protein.

OS Pseudomonas syringae.

FT Key Location/Qualifiers  
FT Domain 45..102  
FT Region 45..79  
FT /label= Hypersensitive\_response\_eliciting\_domain  
FT /label= Acidic\_unit  
FT 79..102  
FT /label= Alpha\_helix

XX WO200198501-A2.

XX 27-DEC-2001.

XX 12-JUN-2001; 2001WO-US018820.

XX 16-JUN-2000; 2000US-021211P.

XX (EDEN-) EDEN BIOSCIENCE CORP.

XX Fan H, Wei Z;

XX WPI; 2002-122282/16.  
XX N-PSDB; AAD27020.

XX New hypersensitive response elicitor proteins comprising spaced apart  
XX domains having an acidic portion linked to an alpha-helix, useful for  
XX imparting disease or stress resistance, controlling insects or enhancing  
XX plant growth.

XX Disclosure; Page 28-29; 99pp; English.

XX The patent discloses hypersensitive response elicitor proteins and  
XX nucleotides encoding such proteins. Hypersensitive response elicitor  
XX proteins comprise an isolated pair or more of spaced apart domains, each  
XX comprising an acidic portion linked to an alpha-helix and capable of  
XX eliciting a hypersensitive response in plants. Sequences of the invention  
XX are used to impart disease resistance to plants, to enhance plant growth,  
XX to control insects and/or to impart stress resistance to plants which  
XX includes resistance to environmental stresses such as climate, air  
XX pollution, chemical and nutritional stress. The method of imparting  
XX disease resistance has the potential for treating previously untreatable  
XX diseases, treating diseases systemically and avoiding the use of  
XX infectious agents or environmentally harmful materials. Hyper- sensitive  
XX response elicitor sequences are used to enhance plant growth which  
XX encompasses greater yield, increased in quantity of seeds produced,  
XX percentage of seeds germinated, plant size and biomass, bigger fruits,  
XX earlier fruit coloration and plant maturation. They are also used for  
XX insect control which encompasses preventing direct insect damage to plant  
XX by feeding injury, interfering with insect larval feeding on the plants,  
XX preventing insects from colonising host plants and releasing phytotoxins.  
XX Sequences of the invention also prevent subsequent disease damage to  
XX plants resulting from insect infection. The present sequence is  
XX Pseudomonas syringae hypersensitive response elicitor protein, dspE

SQ Sequence 424 AA;

Query Match 24.2%; Score 559; DB 5; Length 424;  
Best Local Similarity 36.1%; Pred. No. 2.7e-30;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRQTTEOMQLLAELKSL---LSPOSNAATGAGGNDQTTGVGNAGGLNGRKGTAG 93  
DB 72 KENDSQS--NIAKLIISALIMSLLOMLTNSNKKQDTNQEOPDSQAPFQNNGLG----- 122

QY 94 TTPQSDSQNMLSEMGNNGLDQAITPDGCGGQIGDNPLLKAMLKLIARWMDGSDQFGOP 153  
DB 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151

QY 154 GTGNNSASGTSSSGSPFNLSGGKAPSGNSGNYSPVSTPSTPTSPSTPLDFPS 213  
DB 152 GGDTPATGGGGGGGGTPTATGGG---SGGTPTATGGGEGGVTPTITPOL-----A 200

QY 214 SPTKAAGSTPVTDPHPDVGSGAGIGAGNSVATISAGANTVLHDITTVKAGQVDFGKGT 273  
DB 201 NENRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVDFGHGAT 245

QY 274 FTAGSELGPGGSENGKPLFILEDGLASLKNVTMGDDGADGHIHYG---DAKIDNLHVTN 329  
DB 246 FTADKSMGNGDQENQKPFELAEAGATLKNVLGENEVDGIHVKAQAEVTDINVAQN 305

QY 330 VGEDAITVKPNSAGKSHVEITNSSFEHASDKILQLNADTNLSVDNVKAKDFGTFVRTNG 389  
DB 306 VGEDLITVKGEGGAATNLIKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVRTNG 365

QY 390 GQO-GNWDNLNLSHSAEDGKFSFKSDSEGLNVTSDISLGDEVNHY 435  
DB 366 GKQFDDMSIELNGIEANHGKFAVLRSDSDLLKATGNIAMTDVKHAY 412

RESULT 14

ABB09228  
ID ABB09228 standard; protein; 424 AA.

XX ABB09228;

DT 08-JUL-2002 (first entry)

DE P. syringae hypersensitive response elicitor protein SEQ ID NO:9.

XX Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;  
XX postharvest disease.

OS Pseudomonas syringae.

XX WO200180639-A2.

XX 01-NOV-2001.

XX 17-APR-2001; 2001WO-US012468.

XX 19-APR-2000; 2000US-0198359P.

XX (EDEN-) EDEN BIOSCIENCE CORP.

XX Wei Z, Qiu D, Remick D;

XX WPI; 2002-041357/05.

XX N-PSDB; ABL51713.

XX Inhibiting post harvest disease (caused by Penicillium, Botrytis,  
XX Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a  
XX fruits or vegetables, using hypersensitive response elicitor proteins or  
XX nucleic acids.

XX Disclosure; Page 15-16; 72pp; English.

XX The present invention describes methods for inhibiting post harvest

CC disease or desiccation and enhancing the longevity in a fruits or  
CC vegetables, using hypersensitive response elicitor proteins or  
CC polypeptides or nucleic acids (I) derived from pathogens (e.g. Erwinia  
CC amylovora, E. stewartii, E. chrysanthemi, E. carotovora, Xanthomonas,  
CC Pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter).  
CC (I) has bactericidal activity, and can be used in gene therapy. The  
CC method can be used for inhibiting post harvest disease (caused by  
CC Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and  
CC enhancing the longevity in a fruits or vegetables. The method enables  
CC growers, warehouse packers, shippers and suppliers to process, handle and  
CC store fruit and vegetables with reduced losses caused by post harvest  
CC disease and desiccation, therefore reducing costs to the consumer and  
CC improving quality. The present sequence represents a hypersensitive  
CC response elicitor protein given in the exemplification of the present  
CC invention  
XX  
SQ Sequence 424 AA;  
Query Match 24.2%; Score 559; DB 5; Length 424;  
Best Local Similarity 36.1%; Pred. No. 2.7e-30;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
QY 37 QPIDQTIQMAQLAELIKSL---LSPOSNAATGAGNDQTTGVNAGLNGRKGATG 93  
DB 72 KPNDQS--NIAKLIISALIMSLQMLTNSNKKQDTNQEPDQAPFQNNGLG----- 122  
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGQGGQIGDNPILLKAMKLIARMMDGQSQDFGQP 153  
DB 123 -TPSADS-----GGGG-----TPDATGGG-GTTP-----SATGGG 151  
QY 154 GTGNNSASGTSSSGSGSPFNDLSGGKAPSGNSPNSPVSTFSPSTPTSPSLDPPS 213  
DB 152 GGDTPATGCGGSGGCTPTATGGG---SGGTPATGGGEGGVTPOITPQL-----A 200  
QY 214 SPTKAAGSTPTVDPDPVSGAGIGAGNSVATSGAGNTVLHDTITVKAGVFDGKGT 273  
DB 201 NPNRTSG-----TGVSVDSTAGS-----TEQAGKINNVKDTIKVAGEVFDGHGAT 245  
QY 274 FTAGSELGDSQSENOKPLFILEDGASLKNVTMGDDGADGHIHYG-----DAKIDNLHVTN 329  
DB 246 FTADKSMGNGDQENQKPFELAEAGTLKNVLGENEVDGIHVAKNAQEVTDINVAQN 305  
QY 330 VGEDAITVKPNSAGKSHVEITNSSFEHASDKILOLNADTNLSVDNVKAKDFGTFRNG 389  
DB 306 VGEDLITVKEGGAAVTNLNINSSAKGADDKVQVLNANTHLKIDNFKADDFGTWVRNG 365  
QY 390 GOQ-GNWDNLNLSHSEDGKFSVKSDSEGLNVNTSDISLGDVENHY 435  
DB 366 GKQFDMSIELNGIEANHGKFPALVKSDSDDLKLATGNIAMTQVKHAY 412

RESULT 15  
AAW83014  
ID AAW83014 standard; protein; 197 AA.  
XX  
AC  
XX AAW83014;  
XX  
DT 28-JAN-1999 (first entry)  
XX  
DE Bacillus sp strain KSM-PI5 pectic acid lyase.  
XX  
DE Bacillus sp. strain KSM-PI5; pectic acid lyase; protopectinase;  
KW protopectin; polygalacturonic acid; detergent; cotton; surfactant;  
KW cellulase; protease; bleaching agent.  
XX  
OS Bacillus sp.  
XX  
PN WO9845393-A2.  
XX  
PD 15-OCT-1998.  
XX  
PF 08-APR-1998; 98WO-JP001613.  
XX

PR 09-APR-1997; 97JP-00091142.  
PR 08-SEP-1997; 97JP-00242736.  
XX  
PA (KAOS ) KAO CORP.  
XX  
PI Wada Y, Kasai M, Shikata S, Suzumatsu A, Koike K, Hatada Y;  
PI Kobayashi T, Ito S, Tsumadori M;  
XX  
DR WPI; 1998-568339/48.  
DR N-PSDB; AAV69879.  
XX  
XX Detergent composition containing protopectinase active at alkaline pH -  
PT on protopectin and polygalacturonic acid, provides better removal of  
PT muddy soil.  
XX  
PS Example; Page 71-72; 80pp; English.  
XX  
CC The present invention describes a detergent composition which contains a  
CC protopectinase having an optimum pH 7 or higher against protopectin and  
CC polygalacturonic acid substrates. Inclusion of protopectinase gives a  
CC composition that provides better removal of muddy soil, particularly from  
CC socks. The present sequence represents pectic acid lyase from Bacillus  
CC sp. strain KSM-PI5, which is used in an example from the present  
CC invention. Pectic acid lyase exhibits protopectinase activity  
XX  
SQ Sequence 197 AA;  
Query Match 15.6%; Score 359.5; DB 2; Length 197;  
Best Local Similarity 46.3%; Pred. No. 6.2e-17;  
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;  
QY 253 TVLHDTITVKAGQVDFGKGTFTAG-SELGDSQSENOKPLFILEDGASLKNVTMGDDGA 311  
DB 3 TVVHETIRVPAQTDFGKGTIVVANPNTLGDGQSENOKPIFRLGAGASLKNVWIGAPAA 62  
QY 312 DGIHLYGBAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASDKILOLNADTNL 371  
DB 63 DGVHCYGDCTITNVIEDVGEDALTK--SSGT---VNISGGAAYKAYDKVQINAAGTI 117  
QY 372 SVDNVKAXDFGTFVRTNGGQGNWDLNLSHISAEDGKFSVKSDSE---GLNVNT 423  
DB 118 NIRNFRADDIGKLVKQNGTGYKVMNVENCNISRVKDAILRTDSTSTGRIVNT 172  
Search completed: January 25, 2005, 16:05:58  
Job time : 160 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 15:53:21 ; Search time 40 Seconds  
(without alignments)  
741.104 Million cell updates/aec

Title: US-09-879-248-6  
Perfect score: 2310  
Sequence: 1 MSILTLNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCUS COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	3 US-09-120-927-2	Sequence 2, Appli
2	2310	100.0	447	4 US-09-431-614-6	Sequence 6, Appli
3	559	24.2	424	3 US-09-120-817-2	Sequence 2, Appli
4	559	24.2	424	4 US-09-431-614-14	Sequence 14, Appli
5	359.5	15.6	197	3 US-09-402-668-2	Sequence 2, Appli
6	341	14.8	221	3 US-09-198-956-4	Sequence 4, Appli
7	341	14.8	221	4 US-09-670-141-4	Sequence 4, Appli
8	190.5	8.2	62	3 US-09-402-668-10	Sequence 10, Appli
9	181.5	7.9	2411	3 US-09-268-347-36	Sequence 36, Appli
10	174	7.5	2042	4 US-09-077-098A-6	Sequence 6, Appli
11	165.5	7.2	344	1 US-08-891-254-7	Sequence 7, Appli
12	165.5	7.2	344	4 US-08-819-539-7	Sequence 7, Appli
13	165.5	7.2	344	2 US-09-030-270A-7	Sequence 7, Appli
14	165.5	7.2	344	3 US-08-984-207-7	Sequence 7, Appli
15	165.5	7.2	344	3 US-09-013-587-7	Sequence 7, Appli
16	165.5	7.2	344	4 US-09-086-118-27	Sequence 27, Appli
17	165.5	7.2	344	4 US-09-431-614-15	Sequence 15, Appli
18	165.5	7.2	344	5 PCT-US96-08819-7	Sequence 7, Appli
19	160.5	6.9	907	2 US-09-010-928B-4	Sequence 4, Appli
20	157.5	6.8	2870	4 US-09-479-467A-15	Sequence 15, Appli
21	157.5	6.8	3178	4 US-09-479-467A-4	Sequence 4, Appli
22	156.5	6.8	1912	1 US-08-409-995-4	Sequence 4, Appli
23	156.5	6.8	1912	3 US-08-685-467-4	Sequence 4, Appli
24	156	6.8	2039	4 US-09-077-098A-7	Sequence 7, Appli
25	155.5	6.7	2353	3 US-09-377-155-33	Sequence 33, Appli
26	155.5	6.7	2353	3 US-08-913-942-4	Sequence 4, Appli
27	155.5	6.7	2353	3 US-09-669-974-33	Sequence 33, Appli

28	155.5	6.7	2353	4 US-09-797-862-33	Sequence 33, Appli
29	155.5	6.7	2353	4 US-09-684-707-4	Sequence 4, Appli
30	155.5	6.7	2354	3 US-09-268-347-47	Sequence 47, Appli
31	155	6.7	385	5 PCT-US93-06243-2	Sequence 2, Appli
32	155	6.7	495	2 US-08-794-795-2	Sequence 2, Appli
33	155	6.7	495	3 US-09-249-200-2	Sequence 2, Appli
34	154.5	6.7	975	4 US-09-328-352-4764	Sequence 4764, Ap
35	154	6.7	571	3 US-09-134-001C-3865	Sequence 3865, Ap
36	151	6.5	385	1 US-08-891-254-3	Sequence 3, Appli
37	151	6.5	385	2 US-08-819-539-3	Sequence 3, Appli
38	151	6.5	385	5 PCT-US96-08819-3	Sequence 3, Appli
39	151	6.5	403	2 US-08-200-724A-2	Sequence 2, Appli
40	151	6.5	403	2 US-09-030-270A-3	Sequence 3, Appli
41	151	6.5	403	3 US-08-851-376A-2	Sequence 2, Appli
42	151	6.5	403	3 US-08-984-207-3	Sequence 3, Appli
43	151	6.5	403	3 US-09-013-587-3	Sequence 3, Appli
44	151	6.5	403	4 US-09-086-118-23	Sequence 23, Appli
45	151	6.5	403	4 US-09-431-614-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-120-927-2  
; Sequence 2, Application US/09120927  
; Patent No. 6262018  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jihyun Francis  
; APPLICANT: Beer, Steven V.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; TITLE OF INVENTION: ERWINIA AMYLOVORA AND ITS USE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Nixon, Hartgrave, Devans & Doyle LLP  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,927  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,108  
; FILING DATE: 06-AUG-1977  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1581  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-120-927-2

Query Match 100.0%; Score 2310; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.8e-175;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNTSSPGLFQSGDNLGHHNANSALGOQPIDRQTIEQMAQLLELLKSLLS 60

Db 1 MSILTLNNTSSPGIFQSGDNGLGHHNALSALGQOPIDRQTIQMAQLLAELLSLIS 60  
QY 61 POSGNAATGAGNDQTTGVNAGGLNKRGTAGTTPOSDSQNMLSEMGNGLDQAITPDG 120  
Db 61 POSGNAATGAGNDQTTGVNAGGLNKRGTAGTTPOSDSQNMLSEMGNGLDQAITPDG 120  
QY 121 QCGGQIGDNPPLKAMLKLIARWMDGSDQFGPGTGNNSSASGTSSSGSGSPNDLSGGKA 180  
Db 121 QCGGQIGDNPPLKAMLKLIARWMDGSDQFGPGTGNNSSASGTSSSGSGSPNDLSGGKA 180  
QY 181 PSNGSPSGNYSVPSTFSPSTPTSPPLDFPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240  
Db 181 PSNGSPSGNYSVPSTFSPSTPTSPPLDFPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240  
QY 241 NSVAFTSAGANTVLHDTITVKAGOVFDGKGTFTAGSELGGGSENGKPLFILEDGAS 300  
Db 241 NSVAFTSAGANTVLHDTITVKAGOVFDGKGTFTAGSELGGGSENGKPLFILEDGAS 300  
QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDEDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
Db 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDEDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
QY 361 KILQNLADTNLSVDNVKAKDPTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
Db 361 KILQNLADTNLSVDNVKAKDPTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENVHYKVPMSANLKVAE 447  
Db 421 VNTSDISLGDVENVHYKVPMSANLKVAE 447

RESULT 2

US-09-431-614-6  
; Sequence 6, Application US/09431614  
; Patent No. 6624139  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Schading, Richard L.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS  
; TITLE OF INVENTION: RESISTANCE  
; FILE REFERENCE: 21829/41 (EBC-003)  
; CURRENT APPLICATION NUMBER: US/09/431,614  
; EARLIER FILING DATE: 1999-11-02  
; EARLIER APPLICATION NUMBER: 60/107,243  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver..2.0  
; SEQ ID NO 6  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-431-614-6

Query Match 100.0%; Score 2310; DB 4; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.8e-175;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNTSSPGIFQSGDNGLGHHNALSALGQOPIDRQTIQMAQLLAELLSLIS 60  
Db 1 MSILTLNNTSSPGIFQSGDNGLGHHNALSALGQOPIDRQTIQMAQLLAELLSLIS 60  
QY 61 POSGNAATGAGNDQTTGVNAGGLNKRGTAGTTPOSDSQNMLSEMGNGLDQAITPDG 120  
Db 61 POSGNAATGAGNDQTTGVNAGGLNKRGTAGTTPOSDSQNMLSEMGNGLDQAITPDG 120  
QY 121 QCGGQIGDNPPLKAMLKLIARWMDGSDQFGPGTGNNSSASGTSSSGSGSPNDLSGGKA 180  
Db 121 QCGGQIGDNPPLKAMLKLIARWMDGSDQFGPGTGNNSSASGTSSSGSGSPNDLSGGKA 180  
QY 181 PSNGSPSGNYSVPSTFSPSTPTSPPLDFPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240  
Db 181 PSNGSPSGNYSVPSTFSPSTPTSPPLDFPSSPTKAAGSTPVTDHPDPVGSAGIGAG 240

QY 241 NSVAFTSAGANTVLHDTITVKAGOVFDGKGTFTAGSELGGGSENGKPLFILEDGAS 300  
Db 241 NSVAFTSAGANTVLHDTITVKAGOVFDGKGTFTAGSELGGGSENGKPLFILEDGAS 300  
QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDEDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
Db 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDEDAITVKPNSAGKSKSHVEITNSSFEHSD 360  
QY 361 KILQNLADTNLSVDNVKAKDPTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
Db 361 KILQNLADTNLSVDNVKAKDPTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENVHYKVPMSANLKVAE 447  
Db 421 VNTSDISLGDVENVHYKVPMSANLKVAE 447

RESULT 3

US-09-120-817-2  
; Sequence 2, Application US/09120817  
; Patent No. 6172184  
; GENERAL INFORMATION:  
; APPLICANT: Collmer, Alan  
; APPLICANT: Charkowski, Amy  
; APPLICANT: Alfano, James R.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; TITLE OF INVENTION: PSEUDOMONAS SYRINGAE AND ITS USE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Nixon, Hargrave, Devans & Doyle LLP  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,817  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,107  
; FILING DATE: 06-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1741  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 424 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-120-817-2

Query Match 24.2%; Score 559; DB 3; Length 424;  
Best Local Similarity 36.1%; Pred. No. 2.9e-36;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 OPIDRQTIQMAQLLAELLSLIS---LSPSGNAATGAGNDQTTGVNAGGLNKRGTAG 93  
Db 72 KPNDQS--NIAKULISALIMSILQMLTNSKKQDTNQPDSPQAPFQNNGLG----- 122  
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGQGGQIGDNPPLKAMLKLIARWMDGSDQFGOP 153

Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151  
Qy 154 GTGNASAGTSSGGSPFNLSGKAPSGNSGNYSPVSTPSPSTPTSPSPDPPS 213  
Db 152 GGDTPATGGGGGGGTTATGGG---SGGTPATGGGGGVTPTTQ-----A 200  
Qy 214 SPTKAAGGSTPTVDHDPDVSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVFDKGQT 273  
Db 201 NNRTSG-----TGSVSDTAGS-----TEQAGKINNVKDTIKVGA GEFVDFGHGAT 245  
Qy 274 FTAGSELGCGGQSENGKPLFILEDGASLKNVTMGDDGADGIIHLYG----DAKIDNLHVTN 329  
Db 246 FTADKSMGNGDQENKQPFELAEAGTLKNVNLGENEVGIIHVAKNAQAEVTDINVAQN 305  
Qy 330 VGEDAITVKPNSAGKSHVEITNSSFEHASDKILQNLADTNLSVDNVKAKDFGTGTVRTNG 389  
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADKVVQLNANTHLKIDNFKADDFGTGTVRTNG 365  
Qy 390 GQO-GNWDNLNLSHSAEDGKFSVKSDEGLNVNTSDISLGDVENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDDSLKATGNTAMTDVKHAY 412

## RESULT 4

US-09-431-614-14  
; Sequence 14, Application US/09431614  
; Patent No. 6624139  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Schading, Richard L.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS  
; FILE OF INVENTION: RESISTANCE  
; FILE REFERENCE: 21829/41 (EBC-003)  
; CURRENT APPLICATION NUMBER: US/09/431,614  
; PRIOR FILING DATE: 1999-11-02  
; EARLIER APPLICATION NUMBER: 60/107,243  
; EARLIER FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; TYPE: PRT  
; ORGANISM: Pseudomonas syringae  
US-09-431-614-14

Query Match 24.2%; Score 559; DB 4; Length 424;  
Best Local Similarity 36.1%; Pred. No. 2.9e-36;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
Qy 37 QPIDQTIEQMAQLLAELKSL---LSQSGNNAATGAGNDQTTGVGNAGLNGRKG TAG 93  
Db 72 KPNDQS--NIAKLISALIMSLLOMLTNSGNKQDTNQEQPDSQAPFQNNGLG----- 122  
Qy 94 TTPQSDSQNLMSEMGNGLDQAITPQCGGQIGDNPPLKMLKLIARMQDSQDFGP 153  
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151  
Qy 154 GTGNASAGTSSGGSPFNLSGKAPSGNSGNYSPVSTPSPSTPTSPSPDPPS 213  
Db 152 GGDTPATGGGGGGGTTATGGG---SGGTPATGGGGGVTPTTQ-----A 200  
Qy 214 SPTKAAGGSTPTVDHDPDVSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVFDKGQT 273  
Db 201 NNRTSG-----TGSVSDTAGS-----TEQAGKINNVKDTIKVGA GEFVDFGHGAT 245  
Qy 274 FTAGSELGCGGQSENGKPLFILEDGASLKNVTMGDDGADGIIHLYG----DAKIDNLHVTN 329  
Db 246 FTADKSMGNGDQENKQPFELAEAGTLKNVNLGENEVGIIHVAKNAQAEVTDINVAQN 305  
Qy 330 VGEDAITVKPNSAGKSHVEITNSSFEHASDKILQNLADTNLSVDNVKAKDFGTGTVRTNG 389  
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADKVVQLNANTHLKIDNFKADDFGTGTVRTNG 365

Qy 390 GQO-GNWDNLNLSHSAEDGKFSVKSDEGLNVNTSDISLGDVENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDDSLKATGNTAMTDVKHAY 412

## RESULT 5

US-09-402-668-2  
; Sequence 2, Application US/09402668  
; Patent No. 6172030  
; GENERAL INFORMATION:  
; APPLICANT: WADA, Yasunao  
; APPLICANT: KASAI, Miyuki  
; APPLICANT: SHIKATA, Shitsuw  
; APPLICANT: SUZUMATSU, Atsushi  
; APPLICANT: KOIKE, Kenzo  
; APPLICANT: HATADA, Yuji  
; APPLICANT: KOBAYASHI, Tohru  
; APPLICANT: ITO, Susumu  
; APPLICANT: TSUMADORI, Masaki  
; TITLE OF INVENTION: Detergent Composition  
; FILE REFERENCE: 2173-0116P  
; CURRENT APPLICATION NUMBER: US/09/402,668  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 9-091142 JAPAN  
; PRIOR FILING DATE: 1997-04-09  
; PRIOR APPLICATION NUMBER: 9-242736 JAPAN  
; PRIOR FILING DATE: 1997-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/01613  
; PRIOR FILING DATE: 1998-04-09  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 197  
; TYPE: PRT  
; ORGANISM: Bacillus sp.  
; OTHER INFORMATION: Strain: KSM-P15  
US-09-402-668-2

Query Match 15.6%; Score 359.5; DB 3; Length 197;  
Best Local Similarity 46.3%; Pred. No. 6.6e-21;  
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;  
Qy 253 TVLHDTITVKAGQVFDKGQTTAG-SELGCGGQSENGKPLFILEDGASLKNVTMGDDGA 311  
Db 3 TVVHETIRVPAGQTFDGKQTYVANPNTLGDGSAENQKPIFRLEAGASLKNVVICAPAA 62  
Qy 312 DGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASDKILQNLADTNL 371  
Db 63 DGWHCYGCTTITNVIEDVGEDALTLK--SSGT---VNISGGAAYKAYDKVFOINAAGTI 117  
Qy 372 SYDNVKAQDFGTVRTNGGQGNWDNLNLSHSAEDGKFSVKSDE--GLNVNT 423  
Db 118 NIRNFRAADDIGKLVQNGGTTVYKVMNVENCNISRVKDALTKTDSSTSTGRVNT 172

## RESULT 6

US-09-198-956-4  
; Sequence 4, Application US/09198956  
; Patent No. 6165769  
; GENERAL INFORMATION:  
; APPLICANT: Andersen, Lene N.  
; APPLICANT: Schuelein, Martin  
; APPLICANT: Lange, Niels Erik K.  
; APPLICANT: Bjornvad, Mads E.  
; APPLICANT: Schorr, Kirk  
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus  
; FILE OF INVENTION: Licheniformis  
; FILE REFERENCE: 5377.200-US  
; CURRENT APPLICATION NUMBER: US/09/198,956  
; PRIOR FILING DATE: 1998-11-24  
; EARLIER APPLICATION NUMBER: 1344/97  
; EARLIER FILING DATE: 1997-11-24

EARLIER APPLICATION NUMBER: 60/067,240  
; EARLIER FILING DATE: 1997-12-02  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Bacillus licheniformis  
US-09-198-956-4

Query Match 14.8%; Score 341; DB 3; Length 221;  
Best Local Similarity 41.7%; Pred. No. 2.3e-19;  
Matches 73; Conservative 33; Mismatches 59; Indels 10; Gaps 2;

QY 254 VLHDTITVKGQVDFKGTFTAGSELGDSGQSENGKPLFILEDGASLKNVTMGDDGADG 313  
DB 31 VVHKTIVBKGQTYDGKRLIAGPELGDSGQREDQKPIFKVEDGATLKNVVLGAPADG 90

QY 314 IHLYGDAKIDNLHVTNVGDAITVKNPSAGKSHVEITNSSPEHASDKILOLNADTNLSV 373  
DB 91 VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGSGARLAADKIFQINKASTFTV 145

QY 374 DNVKAKDFGFTVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLNVNTSDISL 428  
DB 146 KNFTADQGGKFIQLGSGSTFKAVVINDNCTITNMKEAIFRTDS-----STSSVTM 195

RESULT 7  
US-09-670-141-4  
; Sequence 4, Application US/09670141  
; Patent No. 6429000  
; GENERAL INFORMATION:  
; APPLICANT: Andersen, Lene N.  
; APPLICANT: Schulein, Martin  
; APPLICANT: Lange, Niels Erik K.  
; APPLICANT: Bjornvad, Mads E.  
; APPLICANT: Schnorr, Kirk  
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus  
; FILE REFERENCE: 5377,200-US  
; CURRENT APPLICATION NUMBER: US/09/670,141  
; CURRENT FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: 09/198,956  
; PRIOR FILING DATE: 1998-11-24  
; PRIOR APPLICATION NUMBER: 1344/97  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/067,240  
; PRIOR FILING DATE: 1997-12-02  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Bacillus licheniformis  
US-09-670-141-4

Query Match 14.8%; Score 341; DB 4; Length 221;  
Best Local Similarity 41.7%; Pred. No. 2.3e-19;  
Matches 73; Conservative 33; Mismatches 59; Indels 10; Gaps 2;

QY 254 VLHDTITVKGQVDFKGTFTAGSELGDSGQSENGKPLFILEDGASLKNVTMGDDGADG 313  
DB 31 VVHKTIVBKGQTYDGKRLIAGPELGDSGQREDQKPIFKVEDGATLKNVVLGAPADG 90

QY 314 IHLYGDAKIDNLHVTNVGDAITVKNPSAGKSHVEITNSSPEHASDKILOLNADTNLSV 373  
DB 91 VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGSGARLAADKIFQINKASTFTV 145

QY 374 DNVKAKDFGFTVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLNVNTSDISL 428  
DB 146 KNFTADQGGKFIQLGSGSTFKAVVINDNCTITNMKEAIFRTDS-----STSSVTM 195

RESULT 8  
US-09-402-668-10  
; Sequence 10, Application US/09402668  
; Patent No. 6172030  
; GENERAL INFORMATION:  
; APPLICANT: WADA, Yasunao  
; APPLICANT: KASAI, Miyuki  
; APPLICANT: SHIKATA, Shitsuw  
; APPLICANT: SUZUMATSU, Atsushi  
; APPLICANT: KOIKE, Kenzo  
; APPLICANT: HATADA, Yuji  
; APPLICANT: KOBAYASHI, Tohru  
; APPLICANT: ITO, Susumu  
; APPLICANT: TSUMADORI, Masaki  
; TITLE OF INVENTION: Detergent Composition  
; FILE REFERENCE: 2173-0116P  
; CURRENT APPLICATION NUMBER: US/09/402,668  
; CURRENT FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 9-091142 JAPAN  
; PRIOR FILING DATE: 1997-04-09  
; PRIOR APPLICATION NUMBER: 9-242736 JAPAN  
; PRIOR FILING DATE: 1997-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/01613  
; PRIOR FILING DATE: 1998-04-09  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 62  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:peptide from  
; OTHER INFORMATION: primer  
US-09-402-668-10

Query Match 8.2%; Score 190.5; DB 3; Length 62;  
Best Local Similarity 65.0%; Pred. No. 3.3e-08;  
Matches 39; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

QY 253 TVLHDTITVKGQVDFKGTFTAG-SELGDSGQSENGKPLFILEDGASLKNVTMGDDGA 311  
DB 3 TVVHETIRVAGQTFDGKQIVVNPNTILGDSQAEKQPIFLRLEAGASLKNVVGAPAA 62

RESULT 9  
US-09-268-347-36  
; Sequence 36, Application US/09268347  
; Patent No. 6335182  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena M.  
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
; FILE REFERENCE: 1038-860  
; CURRENT APPLICATION NUMBER: US/09/268,347  
; CURRENT FILING DATE: 1999-03-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 36  
; LENGTH: 2411  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
US-09-268-347-36

Query Match 7.9%; Score 181.5; DB 3; Length 2411;  
Best Local Similarity 20.6%; Pred. No. 2.6e-05;  
Matches 126; Conservative 64; Mismatches 205; Indels 217; Gaps 26;

QY 4 LTINN-----NTSSSPGLFQSGGDNGLGHN-----ANSALGOPIDROT 43  
DB 1182 IFLANGAAGTDAISNGTISVTGKGLISAGNKEITNVKSALKTYKDTQNTAGATQPA-ANT 1240

QY 44 IEQMAQLLAELKSLLSPOSNAATCAGGNDOTTGVGNAGGLNG-----RKGTAGTTTQ 97  
DB 1241 AEVAKQDLVDLTK----PATGAAGNGADAKAPDTTAAATVGLRGLGWLSAKKTADETQD 1296



```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,254
FILING DATE: 10-JUL-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/475,775
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 14603/10050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-891-254-7

Query Match 7.2%; Score 165.5; DB 1; Length 344;
Best Local Similarity 32.0%; Pred. No. 3.3e-05;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

QY 18 QSGGD---NGLGHNANSALGQPIDROTIEQMAQLLAEL-LKSLSPQSGNAATGAGN 73
DB 134 QPGNDKNGVGGANGAKGAGGQGGGLAEALQIEQLAQLGGGAGAGGAGGAGGAGGA 193
QY 74 DQTGTGNGAGLNGRGKTAGT--TPQSDSQNMLSEMNNGLDQAITPDG--QGGGQIGDN 129
DB 194 DGGGAGGAGGANGADGGNGVNGNQANGPNAGDVNGANGAD-----DGSDDQGLTGV 248
QY 130 PLLKAMKLILARMW-----QSDQFGQPGTGNNSSASGTSSSGSP--FNDLSGGK 179
DB 249 QKLMKILNALVQMVGQGLGGNQAGGSKGAGNAPSGANPGANQPGSADDDQSSGQ 306

RESULT 12
US-08-819-539-7
Sequence 7, Application US/08819539
Patent No. 5859324
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: Hypersensitive Response
TITLE OF INVENTION: Induced Resistance In Plants
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,539
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/475,775
```

```
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 14603/10050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-819-539-7

Query Match 7.2%; Score 165.5; DB 2; Length 344;
Best Local Similarity 32.0%; Pred. No. 3.3e-05;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

QY 18 QSGGD---NGLGHNANSALGQPIDROTIEQMAQLLAEL-LKSLSPQSGNAATGAGN 73
DB 134 QPGNDKNGVGGANGAKGAGGQGGGLAEALQIEQLAQLGGGAGAGGAGGAGGAGGA 193
QY 74 DQTGTGNGAGLNGRGKTAGT--TPQSDSQNMLSEMNNGLDQAITPDG--QGGGQIGDN 129
DB 194 DGGGAGGAGGANGADGGNGVNGNQANGPNAGDVNGANGAD-----DGSDDQGLTGV 248
QY 130 PLLKAMKLILARMW-----QSDQFGQPGTGNNSSASGTSSSGSP--FNDLSGGK 179
DB 249 QKLMKILNALVQMVGQGLGGNQAGGSKGAGNAPSGANPGANQPGSADDDQSSGQ 306

RESULT 13
US-09-030-270A-7
Sequence 7, Application US/09030270A
Patent No. 5977060
GENERAL INFORMATION:
APPLICANT: Zitter, Thomas A.
APPLICANT: Wei, Zhong-Min
TITLE OF INVENTION: INSECT CONTROL WITH A
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,270A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,226
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
```



;  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-030-270A-7

Query Match  
Best Local Similarity 7.2%; Score 165.5; DB 2; Length 344;  
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

QY 18 QSGGD---NGLGHNANSALGQOPIDRQTIEQMAQLLAEI-LKSLSPSGNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKGAGGQGLAEALQIEIQLAQLGGGAGAGGAGGAGG 193

QY 74 DQTTGVNAGGLNGRKGTAGT--TPQSDSQNMLSEMGNNGLDQAITPDG--QGGGGIGDN 129  
Db 194 DGGGAGGAGGAGGAGGAGGQGLAEALQIEIQLAQLGGGAGAGGAGGAGG 193

QY 130 PLLKAWLKLIARMMD-----QSDQFQPGTGNNSASSTSSGGSP--FNDLSGK 179  
Db 249 QKLMKILNALVQMVGGLGGNGQAGGSKGAGNAPASGANPGANPGSADDQSSGQ 306

## RESULT 14

US-08-984-207-7  
; Sequence 7, Application US/08984207  
; Patent No. 6235974  
; GENERAL INFORMATION:  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Beer, Steven V.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED  
; TITLE OF INVENTION: RESISTANCE IN PLANTS BY SEED TREATMENT  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/984,207  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/033,230  
; FILING DATE: 05-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1201  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 344 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-984-207-7

Query Match  
Best Local Similarity 7.2%; Score 165.5; DB 3; Length 344;  
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

QY 18 QSGGD---NGLGHNANSALGQOPIDRQTIEQMAQLLAEI-LKSLSPSGNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKGAGGQGLAEALQIEIQLAQLGGGAGAGGAGGAGG 193

QY 74 DQTTGVNAGGLNGRKGTAGT--TPQSDSQNMLSEMGNNGLDQAITPDG--QGGGGIGDN 129  
Db 194 DGGGAGGAGGAGGAGGAGGQGLAEALQIEIQLAQLGGGAGAGGAGGAGG 193

QY 130 PLLKAWLKLIARMMD-----QSDQFQPGTGNNSASSTSSGGSP--FNDLSGK 179  
Db 249 QKLMKILNALVQMVGGLGGNGQAGGSKGAGNAPASGANPGANPGSADDQSSGQ 306

RESULT 15  
US-09-013-587-7  
; Sequence 7, Application US/09013587  
; Patent No. 6277814  
; GENERAL INFORMATION:  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Beer, Steven V.  
; TITLE OF INVENTION: ENHANCEMENT OF GROWTH IN PLANTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/013,587  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/036,048  
; FILING DATE: 27-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1501  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 344 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-013-587-7

Query Match  
Best Local Similarity 7.2%; Score 165.5; DB 3; Length 344;  
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

QY 18 QSGGD---NGLGHNANSALGQOPIDRQTIEQMAQLLAEI-LKSLSPSGNAATGAGN 73  
Db 134 QPGNDKNGVGGANGAKGAGGQGLAEALQIEIQLAQLGGGAGAGGAGGAGG 193

QY 74 DQTTGVNAGGLNGRKGTAGT--TPQSDSQNMLSEMGNNGLDQAITPDG--QGGGGIGDN 129  
Db 194 DGGGAGGAGGAGGAGGAGGQGLAEALQIEIQLAQLGGGAGAGGAGGAGG 193

QY 130 PLLKAWLKLIARMMD-----QSDQFQPGTGNNSASSTSSGGSP--FNDLSGK 179  
Db 249 QKLMKILNALVQMVGGLGGNGQAGGSKGAGNAPASGANPGANPGSADDQSSGQ 306

Search completed: January 25, 2005, 16:10:54  
Job time : 42 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 16:09:28 ; Search time 149 Seconds  
(without alignments)  
1083.868 Million cell updates/sec

Title: US-09-879-248-6  
Perfect score: 2310  
Sequence: 1 MSILTLNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA.\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTU5\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	9	US-09-835-684-5
2	2310	100.0	447	9	US-09-880-371-5
3	2310	100.0	447	9	US-09-879-248-6
4	2310	100.0	447	14	US-10-010-390-5
5	2310	100.0	447	15	US-10-441-736-6
6	2310	100.0	447	17	US-10-847-142-5
7	559	24.2	424	9	US-09-835-684-9
8	559	24.2	424	9	US-09-880-371-9
9	559	24.2	424	9	US-09-879-248-14
10	559	24.2	424	14	US-10-010-390-9
11	559	24.2	424	15	US-10-441-736-14
12	559	24.2	424	17	US-10-847-142-9
13	256	11.1	276	14	US-10-156-761-13910

14	199.5	8.6	1306	15	US-10-282-122A-64405	Sequence 64405, A
15	193	8.4	694	15	US-10-282-122A-64726	Sequence 64726, A
16	191.5	8.3	584	15	US-10-282-122A-50604	Sequence 50604, A
17	190.5	8.2	226	17	US-10-425-115-329387	Sequence 229387, A
18	188	8.1	639	15	US-10-282-122A-64609	Sequence 64609, A
19	186.5	8.1	591	15	US-10-282-122A-64363	Sequence 64363, A
20	185.5	8.0	1381	15	US-10-282-122A-64895	Sequence 64895, A
21	185	8.0	255	17	US-10-425-115-313677	Sequence 313677, A
22	183	7.9	588	15	US-10-282-122A-64869	Sequence 64869, A
23	182	7.9	606	15	US-10-282-122A-64464	Sequence 64464, A
24	181.5	7.9	484	15	US-09-820-843A-19	Sequence 19, Appl
25	181.5	7.9	484	15	US-10-282-122A-64867	Sequence 64867, A
26	180.5	7.8	1079	10	US-09-820-843A-20	Sequence 20, Appl
27	176.5	7.6	562	15	US-10-282-122A-64514	Sequence 64514, A
28	174	7.5	615	15	US-10-282-122A-64786	Sequence 64786, A
29	174	7.5	2042	14	US-10-192-584-6	Sequence 6, Appl1
30	172.5	7.5	584	15	US-10-282-122A-64903	Sequence 64903, A
31	172	7.4	3507	14	US-10-369-493-5784	Sequence 5784, Ap
32	171.5	7.4	667	15	US-10-282-122A-64494	Sequence 64494, A
33	171	7.4	505	15	US-10-282-122A-62341	Sequence 62341, A
34	169.5	7.3	914	15	US-10-282-122A-64606	Sequence 64606, A
35	169	7.3	525	15	US-10-282-122A-64763	Sequence 64763, A
36	167.5	7.3	1011	15	US-10-282-122A-64589	Sequence 64589, A
37	166.5	7.2	518	15	US-10-282-122A-50634	Sequence 50634, A
38	166.5	7.2	532	15	US-10-282-122A-64658	Sequence 64658, A
39	166	7.2	461	15	US-10-282-122A-64750	Sequence 64750, A
40	166	7.2	603	15	US-10-282-122A-64537	Sequence 64537, A
41	166	7.2	767	15	US-10-282-122A-48384	Sequence 48384, A
42	165.5	7.2	344	9	US-09-086-118-27	Sequence 27, Appl
43	165.5	7.2	344	9	US-09-835-684-11	Sequence 11, Appl
44	165.5	7.2	344	9	US-09-880-371-11	Sequence 11, Appl
45	165.5	7.2	344	9	US-09-879-248-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-835-684-5  
; Sequence 5, Application US/09835684  
; Patent No. US20020019337A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/71  
; CURRENT APPLICATION NUMBER: US/09/835,684  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-835-684-5

Query Match	100.0%;	Score 2310;	DB 9;	Length 447;
Best Local Similarity	100.0%;	Pred. No. 1.4e-151;		
Matches 447;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSILTLNNTSSPGLFQSGDNGILGHNANSALGQPIROTITISMAOLLAEKLSLS	60	
DB	1	MSILTLNNTSSPGLFQSGDNGILGHNANSALGQPIROTITISMAOLLAEKLSLS	60	
QY	61	POSGNAATGAGGNDOTTGVGNAGGLNGRKGTTAGTTPOSDSQNMLSEMGNGLDQAITPDG	120	
DB	61	POSGNAATGAGGNDOTTGVGNAGGLNGRKGTTAGTTPOSDSQNMLSEMGNGLDQAITPDG	120	

QY 121 OGGGOIGDNPLLKAMKLIARMDGSDQFQPGTGNNSSAGTSSSGSPFNDLSGGKA 180  
DB 121 OGGGOIGDNPLLKAMKLIARMDGSDQFQPGTGNNSSAGTSSSGSPFNDLSGGKA 180  
QY 181 PSGNSPGNYSVPSTFSPPTSPSTPLDFPSSPTKAAGGSTPVTDPVGSAGIGAG 240  
DB 181 PSGNSPGNYSVPSTFSPPTSPSTPLDFPSSPTKAAGGSTPVTDPVGSAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGQFTTAGSELGCGQSENOKPLFILEDGAS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGQFTTAGSELGCGQSENOKPLFILEDGAS 300  
QY 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFEHSD 360  
QY 361 KILQNLNADTNLSVDNVKAKDFGTFTVGTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
DB 361 KILQNLNADTNLSVDNVKAKDFGTFTVGTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

## RESULT 2

US-09-880-371-5

; Sequence 5, Application US/09880371

; Patent No. US20020059658A1

; GENERAL INFORMATION:

; APPLICANT: Wei, Zhong-Min

; APPLICANT: DeRoche, Jay

; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC

; TITLE OF INVENTION: PLANTS

; FILE REFERENCE: 21829/91

; CURRENT APPLICATION NUMBER: US/09/880,371

; CURRENT FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: 60/211,585

; PRIOR FILING DATE: 2000-06-15

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 447

; TYPE: PRT

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

DB 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFEHSD 360  
QY 361 KILQNLNADTNLSVDNVKAKDFGTFTVGTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
DB 361 KILQNLNADTNLSVDNVKAKDFGTFTVGTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

## RESULT 3

US-09-879-248-6

; Sequence 6, Application US/09879248

; Patent No. US20020062500A1

; GENERAL INFORMATION:

; APPLICANT: Fan, Hao

; APPLICANT: Wei, Zhong-Min

; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 21829/81

; CURRENT APPLICATION NUMBER: US/09/879,248

; CURRENT FILING DATE: 2001-06-12

; PRIOR APPLICATION NUMBER: 60/212,211

; PRIOR FILING DATE: 2000-06-16

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 447

; TYPE: PRT

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

; ORGANISM: Erwinia amylovora

Query Match 100.0%; Score 2310; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.4e-151;  
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSSPGLFOSGGDNLGCHNANSALGQOPIDROTIQMAQLLAELLKSLLS 60  
DB 1 MSILTLNNNTSSSPGLFOSGGDNLGCHNANSALGQOPIDROTIQMAQLLAELLKSLLS 60  
QY 61 POSGNAATGAGNDQTTGVNAGLNGRGTAGTTPQSDSQNMLSEMNGNGLDQAITPDG 120  
DB 61 POSGNAATGAGNDQTTGVNAGLNGRGTAGTTPQSDSQNMLSEMNGNGLDQAITPDG 120  
QY 121 OGGGOIGDNPLLKAMKLIARMDGSDQFQPGTGNNSSAGTSSSGSPFNDLSGGKA 180  
DB 121 OGGGOIGDNPLLKAMKLIARMDGSDQFQPGTGNNSSAGTSSSGSPFNDLSGGKA 180  
QY 181 PSGNSPGNYSVPSTFSPPTSPSTPLDFPSSPTKAAGGSTPVTDPVGSAGIGAG 240  
DB 181 PSGNSPGNYSVPSTFSPPTSPSTPLDFPSSPTKAAGGSTPVTDPVGSAGIGAG 240  
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGQFTTAGSELGCGQSENOKPLFILEDGAS 300  
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGQFTTAGSELGCGQSENOKPLFILEDGAS 300  
QY 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFEHSD 360  
DB 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFEHSD 360  
QY 361 KILQNLNADTNLSVDNVKAKDFGTFTVGTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
DB 361 KILQNLNADTNLSVDNVKAKDFGTFTVGTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420  
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447  
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

## RESULT 4

US-10-010-390-5

; Sequence 5, Application US/10010390

```
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/248,169
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-010-390-5

Query Match 100.0%; Score 2310; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQPIDROTIEOMAOQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQPIDROTIEOMAOQLLAELLKSLLS 60
QY 61 POSGNAATGAGGNDQTTGVNAGGLNGRKGTTAGTTTQSDSQNNMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGGNDQTTGVNAGGLNGRKGTTAGTTTQSDSQNNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPLLKAMLKLIARWMDGQSDQFGQPGTGNNSSASGTSSTSSGSGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPLLKAMLKLIARWMDGQSDQFGQPGTGNNSSASGTSSTSSGSGSPFNDLSGGKA 180
QY 181 PSNGSPSGNYSPVSTPSTPTSPSLDPFSSPTKAAGGTPVTDHPDPVGSAGIGAG 240
DB 181 PSNGSPSGNYSPVSTPSTPTSPSLDPFSSPTKAAGGTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGOTFTAGSELGQGGQSENOKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGOTFTAGSELGQGGQSENOKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDAITVKPNSAGKXSHVEITNSSFEHSD 360
DB 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDAITVKPNSAGKXSHVEITNSSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFTNGGQGGQNDNLNLSHISAEDGKFSFKVSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFTNGGQGGQNDNLNLSHISAEDGKFSFKVSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/203 (EBC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02

; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/248,169
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6

Query Match 100.0%; Score 2310; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQPIDROTIEOMAOQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQPIDROTIEOMAOQLLAELLKSLLS 60
QY 61 POSGNAATGAGGNDQTTGVNAGGLNGRKGTTAGTTTQSDSQNNMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGGNDQTTGVNAGGLNGRKGTTAGTTTQSDSQNNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPLLKAMLKLIARWMDGQSDQFGQPGTGNNSSASGTSSTSSGSGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPLLKAMLKLIARWMDGQSDQFGQPGTGNNSSASGTSSTSSGSGSPFNDLSGGKA 180
QY 181 PSNGSPSGNYSPVSTPSTPTSPSLDPFSSPTKAAGGTPVTDHPDPVGSAGIGAG 240
DB 181 PSNGSPSGNYSPVSTPSTPTSPSLDPFSSPTKAAGGTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGOTFTAGSELGQGGQSENOKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDKGOTFTAGSELGQGGQSENOKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDAITVKPNSAGKXSHVEITNSSFEHSD 360
DB 301 LKNVTMGDDGADGIHLGYDAKIDNLHVTNVGDAITVKPNSAGKXSHVEITNSSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFTNGGQGGQNDNLNLSHISAEDGKFSFKVSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFTNGGQGGQNDNLNLSHISAEDGKFSFKVSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 6
US-10-847-142-5
; Sequence 5, Application US/10847142
; Publication No. US20040265442A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/197
; CURRENT APPLICATION NUMBER: US/10/847,142
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/835,684
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-847-142-5

Query Match 100.0%; Score 2310; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
```





Db 123 -TPSADS-----CGGG-----TPDATGGG-GDTP-----SATGGG 151  
Qy 154 GTGNNSASGTSSGSPFNDLSGKAPSGNSPSGNYSPVSTFSPSTPTPTSPLDPPS 213  
Db 152 GGDTPATGGGSGGGTPTATGGG---SGGTPTATGGGEGGVTPTQITPOL-----A 200  
Qy 214 SPTKAAGSTPTVDHPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVDPKGQT 273  
Db 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINVVKDTIKVAGAEVFDHGAT 245  
Qy 274 FTAGSELGCGQSENOKPLFILEDGASLKNVTMTGDDGADGIHLYG-----DAKIDNLHVTN 329  
Db 246 FTADKSMGNDGQENQKPMFELAEAGATLKNVNLGENEVDGIHVAKNAQEVITIDNVHAQN 305  
Qy 330 VGEDAITVKPNSAGKSHVEITNSFEHASKILOLNADTNLSVDNVKAKDQFTFVRTNG 389  
Db 306 VGEDLITVKEGGAATVNLNIKSSAKGADDKVVQLNANTHLKIDNFKADDFGTMVRTNG 365  
Qy 390 GOQ-GNWDNLNLSHSAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFPALVKSDSDDLKLTATGNIAMTDVKHAY 412

## RESULT 12

US-10-847-142-9  
; Sequence 9, Application US/10847142  
; Publication No. US20040265442A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/197  
; CURRENT APPLICATION NUMBER: US/10/847,142  
; PRIOR FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: 09/935,684  
; PRIOR FILING DATE: 2001-04-16  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Pseudomonas syringae  
US-10-847-142-9

Query Match 24.2%; Score 559; DB 17; Length 424;  
Best Local Similarity 36.1%; Pred. No. 1.9e-30;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
Qy 37 QPIDRQTEIOMAUQLAELKSL---LSQSGNNAATGAGNDQTTGVGNAGGLNGRKG TAG 93  
Db 72 KPNDQSQ--NIAKLISALIMSLQLMTNSNKKQDTNQEOPDSQAPFQNNGLG-----122  
Qy 94 TTPQSDSONMLSENGNGLDQAITPDQGGGOIGDNPLLKAMLKLTARWMDGSDQFGQP 153  
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151  
Qy 154 GTGNNSASGTSSGSPFNDLSGKAPSGNSPSGNYSPVSTFSPSTPTPTSPLDPPS 213  
Db 152 GGDTPATGGGSGGGTPTATGGG---SGGTPTATGGGEGGVTPTQITPOL-----A 200  
Qy 214 SPTKAAGSTPTVDHPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVDPKGQT 273  
Db 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINVVKDTIKVAGAEVFDHGAT 245  
Qy 274 FTAGSELGCGQSENOKPLFILEDGASLKNVTMTGDDGADGIHLYG-----DAKIDNLHVTN 329  
Db 246 FTADKSMGNDGQENQKPMFELAEAGATLKNVNLGENEVDGIHVAKNAQEVITIDNVHAQN 305

Qy 330 VGEDAITVKPNSAGKSHVEITNSFEHASKILOLNADTNLSVDNVKAKDQFTFVRTNG 389  
Db 306 VGEDLITVKEGGAATVNLNIKSSAKGADDKVVQLNANTHLKIDNFKADDFGTMVRTNG 365  
Qy 390 GOQ-GNWDNLNLSHSAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435  
Db 366 GKQFDDMSIELNGIEANHGKFPALVKSDSDDLKLTATGNIAMTDVKHAY 412

## RESULT 13

US-10-156-761-13910  
; Sequence 13910, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-282  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13910  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13910

Query Match 11.1%; Score 256; DB 14; Length 276;  
Best Local Similarity 31.7%; Pred. No. 1e-09;  
Matches 76; Conservative 35; Mismatches 87; Indels 42; Gaps 7;

Qy 225 VTDHPDPVGS---GIGAGNSVA---FTSAGANQTVLHDT-----ITVKAGQVDPGKGQ 272  
Db 15 VTKRAVIGSAALGLTAGLVTTLSAGAGATSWPEATGSKAVSSIEVSGTYDGKLLK 74  
Qy 273 TTAGSELGCGQSENQKPLFILEDGASLKNVTMTGDDGADGIHLYGDAKIDNLHVTNVE 332  
Db 75 KFGSGDLGTADQSEDQGLPELDGAVLKNVIIGTPAADGVHCLGSCCTLQNVWMLDVG 134  
Qy 333 DAITVKPNSAGKSHVEITNSFEHASKILOLNADTNLSVDNVKAKDQFTFVRTNGGQ 392  
Db 135 DAASFPSKSS--SATYKVIIGGAKSADKVLQFNGAGTLTGTGQVENFGKLVRCGNCK 192  
Qy 393 GNWDNLNLSHSAEDGKFSFKVSDSE-----GLNVNTSDISL-----GDVENHYK 436  
Db 193 TOY-----KPTVVLSDIDATAPGKALVGINSGYGTATLSRIIHGDTKKIK 240

## RESULT 14

US-10-282-122A-64405  
; Sequence 64405, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.



APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 64405  
LENGTH: 1306  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-10-282-122A-64405

Query Match 8.6%; Score 199.5; DB 15; Length 1306;  
Best Local Similarity 26.0%; Pred. No. 5.8e-05;  
Matches 88; Conservative 16; Mismatches 122; Indels 113; Gaps 11;  
QY 8 NNTSSPGLFQSGDNGLNGLGHNANSALGQOPIDROTIEQMAQLLAELLKSLSPSGNAA 67  
Db 754 NGVAGSQPGGAGDGTGGVGGNGRGIDGADGAT----- 789  
QY 68 TCAGGNDOTTVGNAGLNGRGTAGTTQSDSQNMLSEWNGNGLDQAITPDGQGGQIG 127  
Db 790 --AGARGQDGGAGGAGKGRGTGTP-----GGAGPAGTTGSGAGGNG 832  
QY 128 DNPLLKAMLKLIARWMDGQSDQFGOPGTGNNSASGTSSTSSGSPFNDLSGCKA--PSGNSP 186  
Db 833 -----GSGGTGDPDGGNGANGSVFTNNGIGGNGGNGAGPSGAGG 875  
QY 187 SGNYSPVSTPSPPTSPSTPL-----DFPSSPTKAAG-----G 221  
Db 876 SGAG--STFG---ATGSSSIHVNGCGNGGNGDHALSGNAGGNGGSLRG 929  
QY 222 STPVTDHPVGSAGTIGAGNSVAFTSAGANQTVLHDTITVKAQVFDGKQFTTAGSELG 281  
Db 930 SGAGGHGNGGNASRGMGDGTGAGGN-----AGQING-----GAGGNGG 973  
QY 282 DGGQSENQKPLFIL-----EDGASLKNVTMGDDGADG 313  
Db 974 DGGTSGDGNPGAITSGGRGDGGVGGQGGSVAGDGDG 1012

## RESULT 15

US-10-282-122A-64726  
Sequence 64726, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl

APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 64726  
LENGTH: 694  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-10-282-122A-64726  
Query Match 8.4%; Score 193; DB 15; Length 694;  
Best Local Similarity 25.4%; Pred. No. 7.5e-05;  
Matches 86; Conservative 28; Mismatches 140; Indels 84; Gaps 14;  
QY 18 QSGDNGLGGH--NANSALGQOPIDROTIEQMAQLLAELLKSLSPSGNAAATGAGNDQT 76  
Db 256 QTGGDGTGGHGTACTGTGTGGDCAT-----ATAGSKATCGAGDGT 300  
QY 77 TGVGNAGLNGRGTAGTTQSDSQNMLSEWNGNGLDQAITPDGQGGQIGNPLLKAML 136  
Db 301 AAAGGGGNGDGGVA---QGDIAAFGGDGGNGSDGVAAGSGGSGGAGGG---AFV 352  
QY 137 KLIARWMDGQSDQFGOPGTGNNSASGT--SSGSGSPFNDLSGCK--APSGNSPSGNSPV 193  
Db 353 HIATATSTGSGGFG--GNGAASASGADGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 410  
QY 194 STFSPPTSPSTPSPLDFPSSPTKAAGSTPVT--DHPDP-----VGSAG 236  
Db 411 GG-----DGATGSGGSGGNAGIARFDSDPPEAEPDVVGKGGDGGKGGSG 456  
QY 237 IGAGNSVAFTSAGANQTVLHDTITVKAQVFDGKQFTTAGSELGDDGSGSENQKPLFILE 296  
Db 457 LGVGGAGGTGGAGNG-----GAGGLLFGNGN---GGNAGAGG-----D 493  
QY 297 DGASLKNVTMGDDGADGHIHLYGDAKIDNL--HVTNVED 333  
Db 494 GGAGVAGGVGNGGGGGTATTTHEDPVAGVAVGVGGD 531

Search completed: January 25, 2005, 16:22:03  
Job time : 151 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 15:51:17 ; Search time 41 Seconds  
(without alignments)  
1048.997 Million cell updates/sec

Title: US-09-879-248-6  
Perfect score: 2310  
Sequence: 1 MSILTLNNTSSPGLFOSG.....LGDVENHYKVPMSANLKVAE 447  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: piri:\*  
2: piri2:\*  
3: piri3:\*  
4: piri4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	T18447	HrpW protein - Erw
2	353.5	15.3	221	A70045	pectate lyase homo
3	233.5	10.1	219	S68364	pectate lyase (EC
4	207	9.0	1660	A70869	hypothetical glyci
5	199.5	8.6	1306	A70934	hypothetical glyci
6	193	8.4	694	F70868	hypothetical glyci
7	188	8.1	639	D70931	hypothetical glyci
8	186.5	8.1	591	D70523	hypothetical glyci
9	186	8.1	1329	E70917	hypothetical glyci
10	185.5	8.0	1381	E70806	hypothetical glyci
11	185.5	8.0	1489	D70807	hypothetical glyci
12	183.5	7.9	242	A45724	pectate lyase (EC
13	183	7.9	588	F70971	hypothetical glyci
14	182	7.9	606	F70816	hypothetical glyci
15	182	7.9	731	C70974	hypothetical glyci
16	181.5	7.9	484	G70846	hypothetical glyci
17	180.5	7.8	1079	E70807	hypothetical glyci
18	179	7.7	730	F70659	hypothetical glyci
19	179	7.7	1538	F70846	hypothetical prote
20	179	7.7	3016	S77300	hypothetical glyci
21	178	7.7	741	G70917	hypothetical glyci
22	177.5	7.7	1501	F70806	hypothetical glyci
23	177.5	7.7	2232	T34434	hypothetical prote
24	176.5	7.6	562	D70953	hypothetical glyci
25	175	7.6	645	F70825	probable ppe prote
26	174	7.5	615	F70589	hypothetical glyci
27	172.5	7.5	584	G70804	hypothetical glyci
28	172	7.4	3507	T34513	hypothetical prote
29	171.5	7.4	667	A70893	hypothetical glyci

RESULT 1

T18447  
HrpW protein - Erwinia amylovora  
C:Species: Erwinia amylovora  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18447  
R:Gaudriault, S  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z18936  
A:Accession: T18447  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-447 <GAU>  
A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:e1293451; PIDN:CAA741  
A:Experimental source: strain CFBP1430; specific host Pomoideae  
C:Genetics:  
A>Note: hrpW

Query Match	100.0.0%;	Score	2310;	DB	2;	Length	447;					
Best Local Similarity	100.0.0%;	Pred. No.	3.3e-119;									
Matches	447;	Conservative	0;	Mismatches	0;	Indels	0;					
QY	1	MSILTLNNTSSPGLFOSG	DNGLGHNANSALGQQPIDRTQTEQMAQLLAELLKSLLS	60								
Db	1	MSILTLNNTSSPGLFOSG	DNGLGHNANSALGQQPIDRTQTEQMAQLLAELLKSLLS	60								
QY	61	PSGNAATGAGNDQTTG	VGNAGLNGRKGKTAGTTTPOSDSQNMLSEMGNGLDQAITPDG	120								
Db	61	PSGNAATGAGNDQTTG	VGNAGLNGRKGKTAGTTTPOSDSQNMLSEMGNGLDQAITPDG	120								
QY	121	QGGGQIGNPLLKAMLI	ARMMDGSDQFGOPGTGNNSSAGTS	SSGSPFNDLSGGKA	180							
Db	121	QGGGQIGNPLLKAMLI	ARMMDGSDQFGOPGTGNNSSAGTS	SSGSPFNDLSGGKA	180							
QY	181	PSGNSPSGNYSPVST	FPSPPTSP	TSPLD	FPSSPTKAAGG	STPVTDP	PVGSAGTICAG	240				
Db	181	PSGNSPSGNYSPVST	FPSPPTSP	TSPLD	FPSSPTKAAGG	STPVTDP	PVGSAGTICAG	240				
QY	241	NSVAPTSAGANQTVL	HDITITV	KAGQV	FDGKGT	FTAGSEL	DGCGQSE	NKPLFILEDGAS	300			
Db	241	NSVAPTSAGANQTVL	HDITITV	KAGQV	FDGKGT	FTAGSEL	DGCGQSE	NKPLFILEDGAS	300			
QY	301	LKNVTMGDDGADGI	HLVGD	KIDNL	LVTVNG	EDAIT	VKPN	SAGKSHVEIT	NSSFEHSD	360		
Db	301	LKNVTMGDDGADGI	HLVGD	KIDNL	LVTVNG	EDAIT	VKPN	SAGKSHVEIT	NSSFEHSD	360		
QY	361	KILQLNADTNLS	VDNVKAD	FGT	FVRTNGG	QQGN	WDLN	LSHISA	EDKXFS	FVKPS	SEGLN	420
Db	361	KILQLNADTNLS	VDNVKAD	FGT	FVRTNGG	QQGN	WDLN	LSHISA	EDKXFS	FVKPS	SEGLN	420
QY	421	VNTSDISIGD	VENHYKVP	MSANLKV	AE	447						

Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 2  
A70045  
pectate lyase homolog yypA - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: A70045  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C:Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Poulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
ich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A70045  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-221 <KUN>  
A:Cross-references: UNIPROT:Q34310; GB:Z99121; GB:AL009126; NID:G2635827; PIDN:CA815500.  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yypA

Query Match 15.3%; Score 353.5; DB 2; Length 221;  
Best Local Similarity 41.8%; Pred. No. 1e-12;  
Matches 76; Conservative 31; Mismatches 70; Indels 5; Gaps 1;  
QY 250 ANQTVLHDTITVYKAGQVDFKGTFTAGSELGDSQSEKQPLFILEDGSALKNTVMGDD 309  
Db AADKVVHETIIIVPKNTYDGKQRFVAGKELDGDSQSEKQPLFVREDGATLKNVVLGAP 86  
QY 310 GADGHIYLGDAKIDNLHVTNVEDATITVKPNSAGKSHVEITNSSFEHASKILOLNADT 369  
Db AADGVHTYGNVNIQNVKEDVGEDALTVK-----KEGVTIDGSAQKASDKIFQINKAS 141  
QY 370 NLSVDNVKAKDFGTFRVNGGQGGWDLNLSHSAEDGKFSFKSDSEGLNVTSDISLG 429  
Db 142 TFTVKNFTADNGGKFIRQLGSGTFHVDVIIDKCTITNMKEAIFRTDSKTSVTRMTNTRY 201  
QY 430 DV 431  
Db 202 NV 203

RESULT 3  
S68364  
pectate lyase (EC 4.2.2.2) C - fungus (Fusarium solani)  
C:Species: Fusarium solani  
C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S68364  
R:Guo, W.; Gonzalez-Candelas, L.; Kolattukudy, P.E.  
Arch. Biochem. Biophys. 323, 352-360, 1995  
A:Title: Cloning of a new pectate lyase gene pelC from Fusarium solani f. sp. pisi (Nect  
A:Reference number: S68364; MUID:96063610; PMID:7487098  
A:Accession: S68364  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-219 <GUO>  
A:Cross-references: UNIPROT:Q00843; EMBL:U13049  
C:Genetics:  
A:Gene: pelC  
A:Introns: 52/1; 102/3

C:Keywords: carbon-oxygen lyase

Query Match 10.1%; Score 233.5; DB 2; Length 219;  
Best Local Similarity 30.1%; Pred. No. 3.6e-06;  
Matches 63; Conservative 39; Mismatches 72; Indels 35; Gaps 6;  
QY 220 GGSTPVTDPHDPVGSAGIAGNSVAFTSAGANQTVLHDTITVYKAGQVDFKGTFTAGSE 279  
Db 8 GGVPKPTDH-----ISNS-----KVIEKAGQVYDGKWKAYDRSGS 43  
QY 280 LGDG-QGSENGKPLFILEDGSALKNTVMGDDGAGIHIYLGDAKIDNLHVTNVEDATITVK 338  
Db 44 ACKQNEGGDDAVFLLHEGATLKNVIICKQDQSEGVHCKGHTCTLEFVWFEDVCEDAISIK 103  
QY 339 PNASGKSHVEITNSSFEHASKILOLNADTNLSVDNVKAKDFGTFVRT--NGGQGGNWD 396  
Db 104 EDKAGKESW--IIGGAYHSDKVVQHNCGCTVNIINFVEDYKLYRSCGNCCKQCKRN 161  
QY 397 LNLSHISAEDGKFSFKSDSEGLNVTSD 425  
Db 162 VVIEGVTTAKG-----GELAGINANYGD 184

RESULT 4  
A70869  
hypothetical glycine-rich protein Rv2490c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: A70869  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: A70869  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1660 <COL>  
A:Cross-references: UNIPROT:Q53215; GB:AL021246; GB:AL123456; NID:G3261507; PIDN:CAA1606  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: Rv2490c  
C:Superfamily: collagen alpha 1(IV) chain

Query Match 9.0%; Score 207; DB 2; Length 1660;  
Best Local Similarity 28.1%; Pred. No. 0.0012;  
Matches 91; Conservative 19; Mismatches 130; Indels 84; Gaps 15;  
QY 7 NNNTSSPLFQSGGDNGLGHNANSALGQQPIDRQTIEQMAQLLAELKSLSP--QSG 64  
Db 476 NGGAGNGGLV---CDGGAGGGGGAAG-----AGYADMTAIFLGSSGTGEGDG 523  
QY 65 N----AATGAGG-NDQTTGVNAGLNCRKGTAGTTPOSDSQNMLSENGNGLDQA--IT 117  
Db 524 NGGAGGAGGAGAHAGDGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 583  
QY 118 PDGQGGGQIGDNPILKAMLKLIARWMDGSDQFGQPGTGNNSSASGTSSSGG---SPFND 174  
Db 584 GDGGAGGAGGADPAGRA-----GSGVGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 633  
QY 175 LSGGKAPSGNSP-SGNYSPVSTFSPPTPTSTPSPLDFPSPSTKAGGSGTPTVTHDPV 233  
Db 634 GDGGAGGDDGDPGAGKGGAG-----GAGATEGVTGATGATVHSG--G 674  
QY 234 SAGIAGNSVAFTSAGANQTVLHDTITVYKAGQVDFKGTFTAGSELGDSQSENGKPLF 293  
Db 675 NGGKG--GNGADATVAGN-----GKGGAGGNGGLVGGGAG-----710  
QY 294 ILEDG----ASLKNVTMGDDGADG 313  
Db 711 --GDGSGNAGANGANVGEDGADG 732

C/Gene: RV2487C  
 A/Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 8.4%; Score 193; DB 2; Length 694;  
 Best Local Similarity 25.4%; Pred. No. 0.0024;  
 Matches 86; Conservative 28; Mismatches 140; Indels 84; Gaps 14

QY 18 QSGEDNGLGHH-NANSALGQOPIDRTQTEQWALLAELLKLLSPQSNAAATGAGNDQT 76  
 :|||:|||:  
 Db 256 QTGGDGGTGGHTAGTPTGGTGGDAT-----ATAGSKATGGAGGDT 300

QY 77 TGVNAGLNGRKCTAGTTQSDSONMLSEMNGLDQAITPDGCGGQIGDNPILLKAML 136  
 :|||:|||:  
 Db 301 AAAGCGGCGDGGVA---QGDITASAFGDDGNGSDGVAAGSGGSGCGAGG-----APV 352

QY 137 KLITARMMDGSDDFGQPGCTGNNSASSTT-SSSGSGSPFDLSCGK--APSGNSPSGNYSPV 193  
 :|||:|||:  
 Db 353 HIATATSTGSGGFG--GNGAASASADGAGAGGNGGAGGALLFGDGGNGGAGGAGGI 410

QY 194 STSPSPSTPTSPDLFPSPPTKAAGSTPVT--DHPDP-----VGSAG 236  
 :|||:|||:  
 Db 411 GG-----DGATGPGGSGGAGIARFSDPPEAPDVGKGGDGKGGSG 456

QY 237 IGAGNSVAFTSAGANQTVLHDTITVKAGQVDFGKQTFTAGSELGDSQSENQKPLFILE 296  
 :|||:|||:  
 Db 457 LGVGGAGGTGGAGNG-----GAGGLLFGNGN--GGNAGAG-----D 493

QY 297 DGASLKNVTWGDGADGILHYGDADKIDNL-HVTNVGED 333  
 :|||:|||:  
 Db 494 GGAGVAGVGNGGGGGGTATFHEDPVAGVWAVGVGGD 531

RESULT 7  
 D70931  
 hypothetical glycine-rich protein Rv1803c - Mycobacterium tuberculosis (strain  
 C/Species: Mycobacterium tuberculosis  
 C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 31-Mar-2003  
 C/Accession: D70931  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;  
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; H.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.;  
 Nature 393, 537-544, 1998  
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: D70931  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-639 <COL>  
 A;Cross-references: GB:AL022021; GB:AL123456; NID:G3250699; PIDN:CAAI7724.1; P.  
 A;Experimental source: strain H37Rv  
 C/Genetics:  
 A;Gene: Rv1803c  
 C/Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 8.1%; Score 188; DB 2; Length 639;  
 Best Local Similarity 22.9%; Pred. No. 0.004;  
 Matches 96; Conservative 33; Mismatches 144; Indels 146; Gaps 17

QY 7 NNNTSSPGLFQSGDNGLGHNANSALGQOPIDRTQTEQWALLAELLKLLSPQSNAA 66  
 :|||:|||:  
 Db 159 NGNGGSAGLWGSNGGQGGAGANGAAGP-----GK 191

QY 67 ATGAGGNDQT-----TGVGNAGLNGRKGTAGTTQSDSONMLSEMNGLDQAITPDGQ 122  
 :|||:|||:  
 Db 192 AGSGGNGGAGGWYGHGHGGAGGNGNA-TAPCGASAGPDGAGGNG-----GSGGRG 245

QY 123 GQIGDNPILLKAMLKLITARMWDGSDQFGQPGTGNNSASSTSSGSGFPND----- 174  
 :|||:|||:  
 Db 246 GLLFNG-----GNGSVGGMGGQGTNDTAGDSGGLGGNGGAGQGW 290

QY 175 LSGKAPSGNSPSGNYSPVSTFSPSTPTSPDLDFPSSPTKAAGGSTPTVTDHPDPVGS 234  
 :|||:|||:

Db 291 LITNGGQGGDSAGG-----GTDSQTQGVNN-----CASGGSAGIAGNGGDAGL 334  
QY 235 AGIG-----AGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGDDGGQSENQ 289  
Db 335 VNGGAGGNGGAGCASLGT-----IFGSGVGSGGDDGGNGG----- 375  
QY 290 KPLFILEDCASLKNVTMGDDGADGIHLYGDAKIDNLHVTNVGEDALITVKPNSAGKKSHVE 349  
Db 376 ---WLPFGSAS-----CGNGGQG---GDA-----GTNGPAGFGGSAGGGWVG 412  
QY 350 IYNSPEHASDKILQLNADTNLSVDNVKADFGTFVTRVNGGQGN-WDLNLSHISAEDG 407  
Db 413 AVN-----FGPISVQGFELF--GHGDDGNGGDDVGAGSLSTQFG 449  
  
RESULT 8  
B70523  
hypothetical glycine-rich protein Rv0297 - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 31-Mar-2003  
C:Accession: B70523  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: B70523  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-591 <COL>  
A: Cross-references: GB:Z96800; GB:AL123456; NID: g3261800; PIDN: CAB09596.1; PID: e321655;  
A: Experimental source: strain H37Rv  
C: Genetics:  
A: Gene: Rv0297  
C: Superfamily: uncharacterized glycine-rich protein, PE motif containing  
  
Query Match 8.1%; Score 186.5; DB 2; Length 591;  
Best Local Similarity 23.1%; Pred. No. 0.0044; Mismatches 119; Indels 153; Gaps 15;  
Matches 91; Conservative 31;  
  
QY 6 LNNNTSSSPGLFQSGDNGL-----GGHNANSALGQPIDRQTIQMAQLLAELLSLSP 61  
Db 116 ICGANGAPGTQAGSDGGLLFGNGGSGGAPGQ-----AG 152  
QY 62 QSGNAA-----TGAGNDQTTGVNAG----- 83  
Db 153 GAGGAAGFPNGNGGDDGGAGAGGAGGTAGWFFGFGNGGAGGIGVAGINGLGGAGGD 212  
QY 84 -----GLNKRKTAG-----TTPQSDSQNLSEMNGNGLDQAITPDGQGG 123  
Db 213 GGNAGFPNGNGGAGGAGAGVNAVNPGLATPVTAA-----GNGNLNVGVPTAGG 267  
QY 124 GOIGDNPLLKMLKLIARMMDGSDQDFGQFTGNNSASSG-----TSSSGSGSPFNDLSG 177  
Db 268 GADGAN-----GSAIQAG--GAGDGGNASTSGGIGIAQTGGAGGA--GGAGG 312  
QY 178 GMAPSGNSPNSVSPVSTPSTPTSPDPSPTKAGAGSTPVDHDPVGSAGI 237  
Db 313 DGAPGGNGNG-----GSVEHTGATGSSASGGNGATGNGGNGVGAPO- 353  
QY 238 GAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGDDGGQSENQKPLFILED 297  
Db 354 GAGGNGHVSQGSVNT-----AGA--GKGNGGTGGAGGPGGH----- 390  
  
QY 298 GASLKNVTMGDDGADGIHLYGDAKIDNLHVTNVG 331  
Db 391 CGSVLSPGVDSGNGGAGGCGGAGVSATDIAGTG 424

RESULT 9

E70917  
hypothetical glycine-rich protein Rv1450c - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: E70917  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: E70917  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-1329 <COL>  
A: Cross-references: GB:Z95844; GB:AL123456; NID: g3250713; PIDN: CAB09271.1; PID: g2131046  
A: Experimental source: strain H37Rv  
C: Genetics:  
A: Gene: Rv1450c  
C: Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology  
  
Query Match 8.1%; Score 186; DB 2; Length 1329;  
Best Local Similarity 25.2%; Pred. No. 0.013; Mismatches 86; Conservative 25; Indels 104; Gaps 14;  
Matches 86; Conservative 25;  
  
QY 7 NNNNTSSSPGLFQSGDNGLGGHNANSALGQOP-----IDRQTIQMAQLLAELLS 57  
Db 704 NGANATTFG--AKGDDGGHGGFGAGGNGGQGGGPGGLAGNLFQNGIQQVGGSGKGGAG 761  
QY 58 LLSPQSGNAAAT-----GAGNDQTTGVNAGGLN---GRKGTAGTTTQSDSQ 101  
Db 762 GLAGDGGANGNFAFGDNGHGGNGCGAGGQGGGSGAGSTPGAKGAHGFPTSGGD 821  
QY 102 NMLSEMNGLDDQAITPDGQGGQIGDNPPLKMLKLIARMMDGSDQDFGQFTGNNSAS 161  
Db 822 G--GDGNGGNSQVVGNGDGGNGNG-----GSAGTCGNGGR 858  
QY 162 SGTSSSGSPFNDLSGGK-APSGNSPNSVSPVSTPSTPTSPDPSPTKAAK 220  
Db 859 GGDGAFGGSANATNPGENGPN-PCGN-----GGAG 890  
QY 221 GSTPVTDPDHPVGSAGI-----GAGNSVAFTSAGANQTVLHDTITVKA-GQV--PDGKQ 272  
Db 891 GA-----CGAGLNGGAGGNGGLGFGGNGAAGANGVAVGAPGPGGAGGHG 940  
QY 273 TFTAGSELGDDGQSENOKPLFILEDGASLKNVTMGDDGADG 313  
Db 941 AGNGGAGGNGGQG-----VVDGAGGAGGAGDGGGAPG 974  
  
RESULT 10  
E70806  
hypothetical glycine-rich protein Rv3507 - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: E70806  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A: Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: E70806  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-1381 <COL>  
A: Cross-references: GB:AL020222; GB:AL123456; NID: g3261554; PIDN: CAAL1744.1; PID: g292444  
A: Experimental source: strain H37Rv  
C: Genetics:  
A: Gene: Rv3507  
C: Superfamily: collagen alpha 1(IV) chain

Query Match 8.0%; Score 185.5; DB 2; Length 1381;  
Best Local Similarity 25.0%; Pred. No. 0.014;  
Matches 79; Conservative 25; Mismatches 141; Indels 71; Gaps 11;

QY 19 SGDDNGLGCHNANSALGQQ--PIDRTQIBQMAQLLAELKSLSPQSGNAATGAGNDQT 76  
DB 779 SGDDGKGGGSGGSGGSAPI-----GGAGGTGGSGH 814

QY 77 TGVGNAGLNGRKGRTAGTTPQSDSQNMLSEMNGNLDQAITPDGQG--GQIGDNPFLKAM 135  
DB 815 AKGGGAGGI-CAQGTITTVPGNGN--AGDGGNGNAGA---GGNGSGDFGNT----- 863

QY 136 LKLIARMMDQSDQDQCPQTGNNSASSGTSFGSPFNDLGGKAPS--GNSPSGNYSPVS 194  
DB 864 -----TSGASGSGGNGNAGTAGSGAGGTGGTGLSGGNGGNGGNGGDDGGNG 913

QY 195 TFSPPSTPTSPSPIDFPSPPTKAAAGSPFVTDHPDVPVGSAGIGAGNSVAFTSAGANQTV 254  
DB 914 AHGTGVAQVPAISLPTFNGGAGGNGTSGNGAGPAGAPGPTTGGNAGSGGIGDGGN 973

QY 255 LHDITTVKAGQVFDG-----KGQTFTAGS-----ELGDDGQSENOKPLFILED 297  
DB 974 GGD-----CGKGDGADAVNVVFMPTPEQAAITGAGSAGDPTGCGNGGPTGSPMVAPP 1028

QY 298 GASLKNVTMGDDGADG 313  
DB 1029 PTPITQVQQGGDGGAG 1044

RESULT 11  
D70807  
hypothetical glycine-rich protein Rv3514 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: D70807  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: D70807  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1489 <COL>  
A:Cross-references: GB:AL022022; GB:AL123456; NID:93261554; PIDN:CAA17751.1; PID:g292445  
A:Experimental source: strain H37RV  
C:Genetics:  
C:Superfamily: collagen alpha 1(IV) chain

Query Match 8.0%; Score 185.5; DB 2; Length 1489;  
Best Local Similarity 23.9%; Pred. No. 0.015;  
Matches 85; Conservative 26; Mismatches 136; Indels 109; Gaps 12;

QY 10 TSSSPCLFOSGDNGLGCHNANSALGQQPIDRTQIBQMAQLLAELKSLSPQSGNAATG 69  
DB 726 TTGNAGVGGAGSSGAGTNGSGAGG-----TDQGGAGGAG 763

QY 70 AGNDQTTGVG-----NAGLNGRKGRTAGTTPQSDSQNMLSEMNGNLDQAIT 117  
DB 764 GAGADNPTGIGTGGDGTGGAAGAGGAGGAGTGGT-----GGMIGTTGNAGV----- 812

QY 118 PQGGGGQIGDNPFLKAM.LIARMDGQSDQDQCPG--TGNNASGSGTSSGSGSPNDLS 176  
DB 813 --GGAGGQGGDGG-----AGGAGADADQPGATGCTGCFAGGAGGAGGSGSSCA 858

QY 177 GKGAPSGSPSGNYSVPVSTPSTPSPLDPPSPPTKAAGSGTPTVTDHPDVPVGSAG 236  
DB 859 GGTNGSGGA-GGTCCGV-----VAGGAGISFNGSNGGTGGTGGVGGTGGDGGNAG 908

QY 237 TGA-----GNSVAFTSAGAN-----QTVLHDTITVKAG 264  
DB 909 TGAGDPGKGTGGTGGTGGGAGGCGGANFNGTGGTGGKGLNTDGLSATSATG 968

QY 265 QV--PDGKGTPTAGSEL-----GDGQSENOKPLFILEDGASLKNVTMGDDGADG 313  
DB 969 GTGGTGGKGTGGAGDSDSAGTGGTGGAGGAGAGLANTGCTAGNAGIGDGGGQ 1024

RESULT 12  
A45724  
pectate lyase (EC 4.2.2.2) - fungus (Fusarium solani)  
C:Species: Fusarium solani  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A45724  
R:Gonzalez-Candelas, L.; Kolattukudy, P.E.  
J. Bacteriol. 174, 6343-6349, 1992  
A:Title: Isolation and analysis of a novel inducible pectate lyase gene from the phytop  
A:Reference number: A45724; MUID:93015682; PMID:1400187  
A:Accession: A45724  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-242 <GON>  
A:Cross-references: UNIPROT:Q04701; GB:M94691; NID:gl68155; PIDN:AAA33338.1; PID:gl6815  
A:Experimental source: isolate T8  
A:Note: sequence extracted from NCBI backbone (NCBIN:115473, NCBI:115474)  
C:Keywords: carbon-oxygen lyase

Query Match 7.9%; Score 183.5; DB 2; Length 242;  
Best Local Similarity 30.5%; Pred. No. 0.0022;  
Matches 62; Conservative 32; Mismatches 80; Indels 29; Gaps 8;

QY 235 AGIGAGNSVAFT-----SAGANQTVLHDTITVKAGQVFDGKGTPTAGSELGDDGQSENQ 289  
DB 9 AALVGTSSAAVTKTLPKSAGA--TSFTAVPVKGS--YDGMKRFEREKPKVCKGQDETGE 64

QY 290 K-PLFILEDGASLKNVTMGDDGADGIHLGDAKIDNLHVTNVGCEAITVKPSAGKSHV 348  
DB 65 KDAMFILENGATLSNVIIGASQAEVHCCKGTCTLNNVMWADVCEDAVTLKQTSCTS----- 120

QY 349 EITNSSFFHASDKILQNLADNTLSVDNKAQDFGFVFT-----NGGQGNWDLNLSHI 402  
DB 121 YINGGGAFAHSDKILQFNGRGTIVHVKDYAEIDYGLKRSCKGCKNGGPR---NVIVENS 177

QY 403 SAEDKGFVKSDSEGLNVTSD 425  
DB 178 VAVDGGVLC-----GINTNYGD 194

RESULT 13  
F70971  
hypothetical glycine-rich protein Rv3367 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 31-Mar-2003  
C:Accession: F70971  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: F70971  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-588 <COL>  
A:Cross-references: GB:AL009198; GB:AL123456; NID:g3242262; PIDN:CAA15752.1; PID:e12022  
A:Experimental source: strain H37RV  
C:Genetics:  
C:Superfamily: uncharacterized glycine-rich protein, PE motif containing

Query Match 7.9%; Score 183; DB 2; Length 588;

Best Local Similarity 23.1%; Pred. No. 0.0068;  
Matches 83; Conservative 20; Mismatches 151; Indels 106; Gaps 11;  
QY 7 NNNTSSPGLFOSGDNGLGGHN-----ANSALGOQPIDRQTIEQMAQL 50  
DB 208 NCGAGGAAGLGGSGGGGGGGLTGNDCGNPAPVTNPALNAGAGDSNIEPQT----- 260  
QY 51 LAELLKSLSPSGNAATG-----AGNDOTTGVGNA-----GGL 85  
DB 261 -----SVLIGTGGDTGCGAVGNGGAGGAGDANGNPANTSINAGAGGNAAGDGA 315  
QY 86 NGRKGTAGTTPSDSQNMLSEMNGNLDOAITPDGQGG-----GOIGDNPLLKAMLKLIAR 141  
DB 316 NCGAGGAGGQAASAGSSVGGDGGNGGAGGTGTNGHAGGAGGAGGAGGR-----GG 365  
QY 142 MWDGSDQDQGTGNNSSAGTSSSCGSPFNDLGGKAPSGNSPNSGNSVSTSPSPST 201  
DB 366 WLVGNGGNGGNAAGNGAIGGTGGAGGVPA--OQNSALGTQPVGG-----DGGDGGNG 419  
QY 202 PTPSPSLDFPSPPTKAAG-----GSTPVTDPDPVGSAGTGAAGNSVAFTSAGANQTVL 255  
DB 420 GTGTGTCRGDGGSGAGASWLMWNGNGNGGTGGSGGVGNGGIGDGGAGGNATS 479  
QY 256 HDITTVKAGQVDFGKGTFTAGSELGDSQSENQKPLFILEDQASLKNVTMGDDGADGIH 315  
DB 480 TSSIP-----FDAHGNGGAGGDAGHGG-----TGDDGGDGGH 512  
RESULT 14  
H70816  
hypothetical glycine-rich protein Rv0872c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 31-Mar-2003  
C:Accession: H70816  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: C70816  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-606 <COL>  
A: Cross-references: GB:AL022004; GB:AL123456; NID:G3261550; PIDN:CAA17678.1; PID:el25401  
A: Experimental source: strain H37RV  
C: Genetics:  
C: Superfamily: uncharacterized glycine-rich protein, PE motif containing  
Query Match 7.9%; Score 182; DB 2; Length 606;  
Best Local Similarity 26.4%; Pred. No. 0.008;  
Matches 90; Conservative 27; Mismatches 112; Indels 112; Gaps 15;  
QY 16 LFOSGDNGLGGHNANSALGOQPIDRQTIEQMAQLLAELLKSLSPSGNAATGAGNDQ 75  
DB 220 LFGSGAGGCGGNGAAGGVNP-----TPTGTADAGSTGTQ 257  
QY 76 TTGVGNAGLNGRKGTAGTTPSDSQNMLSEMNGNLDOAITPDGQGG-GGQIGDNPLLKA 134  
DB 258 TLG-GNAIGGNGGPGDAGDAMTSGG-----AGSGGNAVSTVNGDVGGEKGK----- 305  
QY 135 MLKLIARMDGSDQDQGTGNNSSAGTSSSGSPFNDLGGKAPS-----GNSPSGN 189  
DB 306 -----GEGAYGAGGAGGSAASIGNAIGG---NGGAGGNAQAPGVGGAGGEG 352  
QY 190 YSPVSTFPSPSTPT-----SPTPLD-FPSSPTKAAGS----- 222  
DB 353 DAQVGTNSFNAEAGNGSGGNGCFDSFASGTCGAGTGGAGGRGGLLIGDGGAGGAGV 412  
QY 223 --TPVTDPDPVGSAGTGAAGNSVAFTSAGANQTVLHDTITVRAGQVFDKG-----QT 273

DB 413 GQTGGSGAPGGGGGAG-GDGAANTDSAGSSR-----KAFGGDGGVGGDGASA 459  
QY 274 FTAGSELGDSQSENQKPLFILEDQASLKNVTMGDDGADGI 314  
DB 460 LQTGGEGIGGGGNG-----GGAG--GLLIGNGGAGGV 490  
RESULT 15  
C70974  
hypothetical glycine-rich protein Rv3388 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 07-Mar-2003  
C:Accession: C70974  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: C70974  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-731 <COL>  
A: Cross-references: GB:AL009198; GB:AL123456; NID:G3242262; PIDN:CAA15773.1; PID:g266166  
A: Experimental source: strain H37RV  
C: Genetics:  
A: Gene: Rv3388  
C: Superfamily: elastin related uncharacterized glycine-rich protein, PE motif containing  
Query Match 7.9%; Score 182; DB 2; Length 731;  
Best Local Similarity 25.3%; Pred. No. 0.01;  
Matches 89; Conservative 26; Mismatches 127; Indels 110; Gaps 15;  
QY 7 NNNTSSPGLFOSGDNGLGGHNANSALGOQPIDRQTIEQMAQLLAELLKSLSPSGNA 66  
DB 399 NNTSAGTGGVGSAGSGTGGNAG-----LIGAGHGG 428  
QY 67 ATGAGNDQTTGVGNAGLNGRKGTAGTTPQ-----SDSQN---MLSEMNGNLDOAIT 117  
DB 429 AGGAGGN-QTGGVGN-GGAGNGGAGGAGGQLYNGGDDGGNGGAGGAGNIAGNGSDGNA 486  
QY 118 PDQGGGGIGDNPLLKAMLKLIARMDGSDQDQGTGNNSSAS-----SQTSSSGSP 171  
DB 487 GHGGAGGS-----ARLI-GAGGHGGDGGAGGNTAGRRADAIAGTGGDGGNG 531  
QY 172 FNDLGGKAPSGNSPNSVSTPSTPTSPSPDFFSPSPTKAAGGSTPTVDHDP 231  
DB 532 GN-----GGLLSGNAGAGHGGGAGGSSTATTG-----TPPTGATGNN----- 570  
QY 232 VGSAGIGAGNSVAFTSAGANQTVLHDTITVRAGQVFDKGQFTTAGSELGDSQSENQKP 291  
DB 571 -GGNG-GAGGTAGTGG-----GIGNGGAGGTGGNAGVA----- 604  
QY 292 LFILEDQASLKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAG 343  
DB 605 LSVGSTGGLGGNGSGGLGGGSLFNGGAGGVGATG-GNGSGIGTPASVG 655  
Search completed: January 25, 2005, 16:10:17  
Job time : 52 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 15:23:52 ; Search time 198 Seconds  
(without alignments)  
1298.952 Million cell updates/sec

Title: US-09-879-248-6  
Perfect score: 2310  
Sequence: 1 MSILTLNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 02:\*

1: uniprot\_sprot: \*  
2: uniprot\_treml: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	054508	O54508 erwinia amy
2	2277	98.6	447	Q9LAW2	Q9LAW2 erwinia amy
3	1916.5	83.0	450	Q6XDB8	Q6XDB8 erwinia pyr
4	1916.5	83.0	450	AAQ17046	AAQ17046 erwinia p
5	1916.5	83.0	450	AAS45453	AAS45453 erwinia p
6	1005	43.5	479	Q6RK52	Q6RK52 pectobacter
7	1005	43.5	479	Q6D5C8	Q6D5C8 erwinia car
8	1005	43.5	479	AAS20352	AAS20352 pectobact
9	705	30.5	138	Q79AW7	Q79AW7 erwinia amy
10	589	25.5	424	Q8RP12	Q8RP12 pseudomonas
11	567.5	24.6	441	Q87327	Q87327 pseudomonas
12	559	24.2	424	Q87264	Q87264 pseudomonas
13	559	24.2	424	Q7C415	Q7C415 pseudomonas
14	472.5	20.5	386	Q849P3	Q849P3 pseudomonas
15	423.5	18.3	307	Q8KKZ2	Q8KKZ2 rhizobium e
16	415.5	18.0	380	Q8XVQ5	Q8XVQ5 raletonia s
17	362.5	15.7	224	Q8RHW0	Q8RHW0 bacillus sp
18	354	15.3	203	Q8PIH5	Q8PIH5 xanthomonas
19	353.5	15.3	221	Q34310	Q34310 bacillus su
20	349.5	15.1	222	Q9X6Z2	Q9X6Z2 bacillus sp
21	346.5	15.0	324	Q8PBA4	Q8PBA4 xanthomonas
22	341	14.8	222	Q8LOR5	Q8LOR5 bacillus sp
23	256	11.1	276	Q829M8	Q829M8 streptomyce
24	226.5	9.8	266	Q9EX16	Q9EX16 streptomyce
25	216.5	9.4	215	Q00843	Q00843 nectria hae
26	207	9.0	1660	Q79FD4	Q79FD4 mycobacteri
27	207	9.0	1660	CAB55496	CAB55496 mycobacte
28	207	9.0	1665	Q7B721	Q7B721 mycobacteri
29	206	8.9	252	Q7S4B7	Q7S4B7 neurospori
30	199.5	8.6	1306	Q7D9L6	Q7D9L6 mycobacteri
31	199.5	8.6	1306	Q7UIQ7	Q7UIQ7 mycobacteri

32 199.5 8.6 1306 2 Q6MX28  
33 199.5 8.6 1306 2 CAE55300  
34 194.5 8.4 233 2 Q00845  
35 194 8.4 1928 2 Q8T9H1  
36 193 8.4 240 2 O93877  
37 193 8.4 694 2 Q7D724  
38 193 8.4 694 2 CAE55495  
39 191 8.3 555 2 Q8SX80  
40 191 8.3 929 2 Q9NBL3  
41 191 8.3 929 2 Q9NBW0  
42 191 8.3 929 2 AAS64813  
43 191 8.3 939 2 Q7KRH9  
44 191 8.3 939 2 Q9NHQ0  
45 191 8.3 939 2 AAS64812

#### ALIGNMENTS

RESULT 1  
O54508  
ID O54508 PRELIMINARY; PRT; 447 AA.  
AC O54508;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
DE HrpW protein.  
GN Name=hrpW;  
OS Erwinia amylovora  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Erwinia.  
OX NCBI\_TaxID=552;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CFBP1430;  
RX MEDLINE=98086111; PubMed=9426142;  
RA Gaudriault S., Malandrin L., Paulin J.P., Barny M.A.;  
RT "DspA, an essential pathogenicity factor of Erwinia amylovora showing  
homology with AvrE of Pseudomonas syringae, is secreted via the Hrp  
secretion pathway in a DspB-dependent way.";  
RL Mol. Microbiol. 26:1057-1069(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CFBP1430;  
RX MEDLINE=98316710; PubMed=9654138;  
RA Gaudriault S., Brisset M.N., Barny M.A.;  
RT "HrpW of Erwinia amylovora, a new Hrp-secreted protein.";  
RL FEBS Lett. 428:224-228(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ea321;  
RA Kim J.F., Zumboff C.H., Beer S.V.;  
RT "HrpW, a new harpin of Erwinia amylovora, is a member of a family of  
pectate lyases.";  
RL Phytopathology 87:0-0(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ea321;  
RX MEDLINE=98422475; PubMed=9748455;  
RA Kim J.F., Beer S.V.;  
RT "HrpW of Erwinia amylovora, a new harpin that contains a domain  
homologous to pectate lyases of a distinct class.";  
RL J. Bacteriol. 180:5203-5210(1998).  
DR EMBL; Y13831; CAA74158.1; -;  
DR EMBL; U94513; AAC62314.1; -;  
DR FIR; T18447; T18447.  
DR HSP; Q9RHW0; 1EE6.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0030570; F:pectate lyase activity; IEA.  
DR InterPro; IPR011050; Pectin\_lyase\_like.  
DR InterPro; IPR004898; Pect\_lyase.  
DR Pfam; PF03211; Pectate\_lyase; 1.  
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Qy	61	POSGNAATGAGNDQTTG	VGNAGLNGRKG	TAGTTT	QSDSQNMLSEMNGNGLDQA	ITPDG	120								
Db	61	POSGNAATGAGNDQTTG	VGNAGLNGRKG	TAGTTT	QSDSQNMLSEMNGNGLDQA	ITPDG	120								
Qy	121	QGGQIGDNP	LLKAMLIAR	MDGSDQ	DPGQGTGNNSSASG	TSSGSPFNDLSGGKA	180								
Db	121	QGGQIGDNP	LLKAMLIAR	MDGSDQ	DPGQGTGNNSSASG	TSSGSPFNDLSGGKA	180								
Qy	181	PSGNSP	SGNYSVST	SPSTPT	SPSPPLDPP	SPSPTKAAGSTP	VTDPVGSAGIGAG	240							
Db	181	PSGNSP	SGNYSVST	SPSTPT	SPSPPLDPP	SPSPTKAAGSTP	VTDPVGSAGIGAG	240							
Qy	241	NSVFTS	SAGANQ	TVLHDT	ITTVKAGQ	VFDGKQFT	TAGSELG	DGQSGNQKPLFILEDGAS	300						
Db	241	NSVFTS	SAGANQ	TVLHDT	ITTVKASQ	VFDGKQFT	TAGSELG	DGQSGNQKPLFILEDGAS	300						
Qy	301	LKNVTMG	DDGADG	IHLYGDAK	IDLNLHVTNVGED	AITVKPNS	SAGKSHVEITNSSFEH	ASD	360						
Db	301	LKNVTMG	DDGADG	IHLYGDAK	IDLNLHVTNVGED	AITVKPNS	SAGKSHVEITNSSFEH	ASD	360						
Qy	361	KILOLN	ADTNLS	VDNVKAKD	FGT	FVRTNGGQ	GNWDLNLSHISAEDG	KFSFVKSDSEGLN	420						
Db	361	KILOLN	ADTNLS	VDNVKAKD	FGT	FVRTNGGQ	GNWDLNLSHISAEDG	KFSFVKSDSEGLN	420						
Qy	421	VNTSD	ISLSD	VENHYK	VPMSANL	KVAE	447								
Db	421	VNTSD	ISLSD	VENHYK	VPMSANL	KVAE	447								
RESULT 3															
ID	Q6XDB8	PRELIMINARY; PRT; 450 AA.													
AC	Q6XDB8;														
DT	05-JUL-2004	(Tremblrel. 27, Created)													
DT	05-JUL-2004	(Tremblrel. 27, Last sequence update)													
DT	05-JUL-2004	(Tremblrel. 27, Last annotation update)													
DE	HrpW.														
GN	Name=hrpW;														
OS	Erwinia pyrifoliae.														
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;														
OC	Enterobacteriaceae; Erwinia.														
OX	NCBI_TaxID=79967;														
RN	[1]														
RP	SEQUENCE FROM N.A.														
RC	STRAIN=WT3;														
RA	Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,														
RA	Hur J.H., Lim C.K.;														
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.														
RN	[2]														
RP	SEQUENCE FROM N.A.														
RC	STRAIN=WT3;														
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,														
RA	Hur J.H., Lim C.K.;														
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.														
DR	EMBL; AY237642; AAQ17046.1; -														
DR	EMBL; AY530755; AAS45453.1; -														
DR	InterPro; IPR011050; Pectin_lyas_like.														
DR	InterPro; IPR004898; Pect_lyase.														
DR	Pfam; PF03211; Pectate_lyase; 1.														
SQ	SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;														
Query Match 83.0%; Score 1916.5; DB 2; Length 450;															
Best Local Similarity 83.2%; Pred. No. 2.9e-90;															
Matches 376; Conservative 23; Mismatches 46; Indels 7; Gaps 2;															
Qy	1	MSILTLNNNTSS	PGLFQSG	DNGLG	GHNANS	ALGQ	QPID	QRTIE	QMAQL	LAELLK	SKLLS	60			
Db	1	MSVLT	LNISIP	SSQGLF	KPGD	NGLSG	QNTNS	AQGH	PI	DR	QRTIE	QMAQL	LAELLK	PLLS	60
Qy	61	POSGNAATGAGNDQTTG	VGNAGLNGRKG	TAGTTT	QSDSQNMLSEMNGNGLDQA	ITPDG	120								
Db	61	POADNAA--	AGSND	QTN	GVGNAG	GLTG	QNGA	AGTT	FQ	SV	NTQ	MLSEMNGNGLDQA	ITPDG	118	

Qy	61	POSGNAATGAGNDQTTG	VGNAGLNGRKG	TAGTTT	QSDSQNMLSEMNGNGLDQA	ITPDG	120							
Db	61	POSGNAATGAGNDQTTG	VGNAGLNGRKG	TAGTTT	QSDSQNMLSEMNGNGLDQA	ITPDG	120							
Qy	121	QGGQIGDNP	LLKAMLIAR	MDGSDQ	DPGQGTGNNSSASG	TSSGSPFNDLSGGKA	180							
Db	121	QGGQIGDNP	LLKAMLIAR	MDGSDQ	DPGQGTGNNSSASG	TSSGSPFNDLSGGKA	180							
Qy	181	PSGNSP	SGNYSVST	SPSTPT	SPSPPLDPP	SPSPTKAAGSTP	VTDPVGSAGIGAG	240						
Db	181	PSGNSP	SGNYSVST	SPSTPT	SPSPPLDPP	SPSPTKAAGSTP	VTDPVGSAGIGAG	240						
Qy	241	NSVFTS	SAGANQ	TVLHDT	ITTVKAGQ	VFDGKQFT	TAGSELG	DGQSGNQKPLFILEDGAS	300					
Db	241	NSVFTS	SAGANQ	TVLHDT	ITTVKASQ	VFDGKQFT	TAGSELG	DGQSGNQKPLFILEDGAS	300					
Qy	301	LKNVTMG	DDGADG	IHLYGDAK	IDLNLHVTNVGED	AITVKPNS	SAGKSHVEITNSSFEH	ASD	360					
Db	301	LKNVTMG	DDGADG	IHLYGDAK	IDLNLHVTNVGED	AITVKPNS	SAGKSHVEITNSSFEH	ASD	360					
Qy	361	KILOLN	ADTNLS	VDNVKAKD	FGT	FVRTNGGQ	GNWDLNLSHISAEDG	KFSFVKSDSEGLN	420					
Db	361	KILOLN	ADTNLS	VDNVKAKD	FGT	FVRTNGGQ	GNWDLNLSHISAEDG	KFSFVKSDSEGLN	420					
Qy	421	VNTSD	ISLSD	VENHYK	VPMSANL	KVAE	447							
Db	421	VNTSD	ISLSD	VENHYK	VPMSANL	KVAE	447							
RESULT 2														
ID	Q9LAW2	PRELIMINARY; PRT; 447 AA.												
AC	Q9LAW2;													
DT	01-OCT-2000	(Tremblrel. 15, Created)												
DT	01-OCT-2000	(Tremblrel. 15, Last sequence update)												
DT	01-JUN-2003	(Tremblrel. 24, Last annotation update)												
DE	Harpin HrpW.													
GN	Name=hrpW;													
OS	Erwinia amylovora.													
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;													
OC	Enterobacteriaceae; Erwinia.													
OX	NCBI_TaxID=552;													
RN	[1]													
RP	SEQUENCE FROM N.A.													
RC	STRAIN=Ea246;													
RA	Kim J.F., Laby R.J., Beer S.V.;													
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.													
DR	EMBL; AF083620; AAF63402.1; -													
DR	HSP; Q9RHW0; IE66.													
DR	GO:0005576; C:extracellular; IEA.													
DR	GO:0030570; F:pectate lyase activity; IEA.													
DR	InterPro; IPR011050; Pectin_lyas_like.													
DR	InterPro; IPR004898; Pect_lyase.													
DR	Pfam; PF03211; Pectate_lyase; 1.													
SQ	SEQUENCE 447 AA; 45340 MW; 0BBAAE3871EDC2F6 CRC64;													
Query Match 98.6%; Score 2277; DB 2; Length 447;														
Best Local Similarity 98.7%; Pred. No. 1.1e-108;														
Matches 441; Conservative 3; Mismatches 3; Indels 0; Gaps 0;														
Qy	1	MSILTLNNNTSS	PGLFQSG	DNGLG	GHNANS	ALGQ	QPID	QRTIE	QMAQL	LAELLK	SKLLS	60		
Db	1	MSILTLNS	TSSPGLF	QSGD	NGLG	HNANS	ALGQ	PI	DR	QRTIE	QMAQL	LAELLK	SKLLS	60

```

QY 121 QGGGQIGDNPPLKAMKLIARMDGSDQDQFGPCTGNNSASCTSSGGSPFNDLS- --- 176
DB 119 QGGGQIGDNPPLKAMKLIARMDGSDQDQFGPCTGNNSASCTSSGGSPFNDLS- --- 178
QY 177 -GKAPSGNSPSGNSPVSTFSPPTPTSPSLDPPSPPTKAAGSGTPTVTHDPPVGS 235
DB 179 LGSSSLGKASSGGTSPSTNSFSPPTPTSPSLDPPSPPTSCAGSGTPTVTHDPPVGS 238
QY 236 GIGAGNSVAFSTAGANOTVLHDTITVKAGQVDFGKGTQTTAGSELGGGSGENKPLFIL 295
DB 239 GVGAGNSVGFSTAGANPTVLHDTIIVKAGQEFDFGKGTQTTAGSELGGGSGENKPLFIL 298
QY 296 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 355
DB 299 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 358
QY 356 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 415
DB 359 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 418
QY 416 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 447
DB 419 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 450

RESULT 4
AAQ17046
ID AAQ17046 PRELIMINARY; PRT; 450 AA.
AC AAQ17046;
DT 03-MAR-2004 (Tremblrel. 27, Created)
DT 03-MAR-2004 (Tremblrel. 27, Last sequence update)
DE HRPW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shreetha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RT "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RT and expression of hrpB encoding elicitor of the hypersensitive
RT response."
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 83.0%; Score 1916.5; DB 2; Length 450;
Best Local Similarity 83.2%; Pred. No. 2.9e-90;
Matches 376; Conservative 23; Mismatches 46; Indels 7; Gaps 2;

QY 1 MSLLTLNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQMAQLLAEKLSILS 60
DB 1 MSVLTINISIPSSQGLFKPEDNGLSGQNTNSAQGHPIRQTIQMAQLLAEKLSILS 60
QY 61 PQSGNAATGAGGNDQTTGVNAGGLNGRGTAGTTTQSDSQNNLSMGNGLDQAITPDG 120
DB 61 PQADNAA--AGSNDQTNVGNAGGLTGQGAAGTTTQSVNQTMLGEMGNGLDQAITPDG 118
QY 121 QGGGQIGDNPPLKAMKLIARMDGSDQDQFGPCTGNNSASCTSSGGSPFNDLS- --- 176
DB 119 QGGGQIGDNPPLKAMKLIARMDGSDQDQFGPCTGNNSASCTSSGGSPFNDLS- --- 178
QY 177 -GKAPSGNSPSGNSPVSTFSPPTPTSPSLDPPSPPTKAAGSGTPTVTHDPPVGS 235
DB 179 LGSSSLGKASSGGTSPSTNSFSPPTPTSPSLDPPSPPTSCAGSGTPTVTHDPPVGS 238
QY 236 GIGAGNSVAFSTAGANOTVLHDTITVKAGQVDFGKGTQTTAGSELGGGSGENKPLFIL 295
DB 239 GVGAGNSVGFSTAGANPTVLHDTIIVKAGQEFDFGKGTQTTAGSELGGGSGENKPLFIL 298
QY 296 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 355
DB 299 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 358
QY 356 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 415
DB 359 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 418
QY 416 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 447
DB 419 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 450

```

```

QY 296 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 355
DB 299 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 358
QY 356 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 415
DB 359 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 418
QY 416 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 447
DB 419 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 450

RESULT 5
AAS45453
ID AAS45453 PRELIMINARY; PRT; 450 AA.
AC AAS45453;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DE HRPW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shreetha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RT "Identification of hrp genes cluster and characterization of HR
RT elicitor hrpNEP gene in Erwinia pyrifoliae."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY530755; AAS45453.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 83.0%; Score 1916.5; DB 2; Length 450;
Best Local Similarity 83.2%; Pred. No. 2.9e-90;
Matches 376; Conservative 23; Mismatches 46; Indels 7; Gaps 2;

QY 1 MSLLTLNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQMAQLLAEKLSILS 60
DB 1 MSVLTINISIPSSQGLFKPEDNGLSGQNTNSAQGHPIRQTIQMAQLLAEKLSILS 60
QY 61 PQSGNAATGAGGNDQTTGVNAGGLNGRGTAGTTTQSDSQNNLSMGNGLDQAITPDG 120
DB 61 PQADNAA--AGSNDQTNVGNAGGLTGQGAAGTTTQSVNQTMLGEMGNGLDQAITPDG 118
QY 121 QGGGQIGDNPPLKAMKLIARMDGSDQDQFGPCTGNNSASCTSSGGSPFNDLS- --- 176
DB 119 QGGGQIGDNPPLKAMKLIARMDGSDQDQFGPCTGNNSASCTSSGGSPFNDLS- --- 178
QY 177 -GKAPSGNSPSGNSPVSTFSPPTPTSPSLDPPSPPTKAAGSGTPTVTHDPPVGS 235
DB 179 LGSSSLGKASSGGTSPSTNSFSPPTPTSPSLDPPSPPTSCAGSGTPTVTHDPPVGS 238
QY 236 GIGAGNSVAFSTAGANOTVLHDTITVKAGQVDFGKGTQTTAGSELGGGSGENKPLFIL 295
DB 239 GVGAGNSVGFSTAGANPTVLHDTIIVKAGQEFDFGKGTQTTAGSELGGGSGENKPLFIL 298
QY 296 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 355
DB 299 EDGASLKNTVMGDDGADGHIHYGDAKIDNLHVTNVGEDAITVKPNSAGKSHVITNSSP 358
QY 356 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 415
DB 359 EHASDKILOLNADTNLSVDNVKAKDFGTFRVINGGQGNWDLNLSHISAEDGKFSFVKSD 418
QY 416 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 447
DB 419 SEGLNVNTSDISLGVNENHYKVPMSANLVAE 450

```

OS Erwinia carotovora subsp. atroseptica SCRI1043.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_TaxID=218491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SCRI1043;  
RA Bell K.S., Sebahia M., Pritchard L., Holden M., Hyman L.J.,  
RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,  
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,  
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,  
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,  
RA Almond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;  
RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
RL EMBL: BX950851; CAG75014.1;  
DR EMBL: BX950851; CAG75014.1;  
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;  
  
Query Match 43.5%; Score 1005; DB 2; Length 479;  
Best Local Similarity 49.9%; Pred. No. 1.1e-43;  
Matches 231; Conservative 60; Mismatches 102; Indels 70; Gaps 15;  
  
QY 8 NNTSSPGLFQSGGDNGLGHNANSALGQOPIDRQITQMAQLLAELLSL----- 59  
Db 61 NTPSSTDG--SSQQAQI-GNGSSALGQN-----GSAADMLKLEALIPSKNGQEA 110  
QY 60 -SQSGNAATGAGN---DQTTGVNAGLNGRKGTTAGTTTQSDSQNMLSEMNGNLDQA 115  
Db 111 GNPLSSGSSGAAGNCGASPLTSGSGAGVGGAQ-----NPEDLSRLSQDSAGSALNNA 165  
QY 116 ITPDQGGGGIGDNPLKMLKLIARMQGSQDFQPGTGNNSASGTSSSGSGSPFNDL 175  
Db 166 INPTADGGGQLSGNDLLKALLELIGNLMDSKQGEFGQP-----QSGSQSGGSP--- 215  
QY 176 SGKAPSGNSPNSGVSVSTFSPSTPTSPSLDFPSSPTKAAG-----STPVTDPD 230  
Db 216 STG-APQASSGGGGSPAA-----PSAPSSVGGNGGAASAPLTAAPT 257  
QY 231 PV--GSAG-----IGAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGD 282  
Db 258 GVDGSSAASPASTAGAG-PVSPPTASANTVNDTIKVGGEVFDGKGTFTAGSELGD 316  
QY 283 GGQSENOKPLFILEDGASLKNVTMGDDGADGIHLGYDAKIDNLHVTNVGEDALTVPNSA 342  
Db 317 GGQAEQKPLPELAQAGATLKNVFGDNDADGVHVRGDAKIDNVHVTNVGEDALTVPNS- 375  
QY 343 GKSHVEITNSSEFHASDKILQNLADTNLSVDNVKAKDFGTFVTRTNGGQGGNDLNLSHI 402  
Db 376 GKPAVETITNSAQASDKIFQLNADANLTIDNFKAKDFGTFVTRTNGGQGGNDLNLNSI 435  
QY 403 SAEDGKFSFKVSDSEGLNVNTSDISLGDVENHYKVPMSANLKV 445  
Db 436 DAENGKFSFKVSDSEGLNVKNINLTNNVNNHYKVPDSANLQV 478  
  
RESULT 8  
AAS20352 PRELIMINARY; PRT; 479 AA.  
ID AAS20352  
AC AAS20352;  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE HRPW.  
GN HRPW.  
OS Pectobacterium atrosepticum.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_TaxID=29471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SCRI1039;  
RA Holvea M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,  
RA Birch P.R.J., Toth I.K.;  
RA Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AV496066; AAS20352.1;  
DR InterPro: IPR011050; Pectin\_lyase\_like.  
DR InterPro: IPR004898; Pect\_lyase.  
DR Pfam: PF03211; Pectate\_lyase; 1.  
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;  
  
Query Match 43.5%; Score 1005; DB 2; Length 479;  
Best Local Similarity 49.9%; Pred. No. 1.1e-43;  
Matches 231; Conservative 60; Mismatches 102; Indels 70; Gaps 15;  
  
QY 8 NNTSSPGLFQSGGDNGLGHNANSALGQOPIDRQITQMAQLLAELLSL----- 59  
Db 61 NTPSSTDG--SSQQAQI-GNGSSALGQN-----GSAADMLKLEALIPSKNGQEA 110  
QY 60 -SQSGNAATGAGN---DQTTGVNAGLNGRKGTTAGTTTQSDSQNMLSEMNGNLDQA 115  
Db 111 GNPLSSGSSGAAGNCGASPLTSGSGAGVGGAQ-----NPEDLSRLSQDSAGSALNNA 165  
QY 116 ITPDQGGGGIGDNPLKMLKLIARMQGSQDFQPGTGNNSASGTSSSGSGSPFNDL 175  
Db 166 INPTADGGGQLSGNDLLKALLELIGNLMDSKQGEFGQP-----QSGSQSGGSP--- 215  
QY 176 SGKAPSGNSPNSGVSVSTFSPSTPTSPSLDFPSSPTKAAG-----STPVTDPD 230  
Db 216 STG-APQASSGGGGSPAA-----PSAPSSVGGNGGAASAPLTAAPT 257  
QY 231 PV--GSAG-----IGAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGD 282  
Db 258 GVDGSSAASPASTAGAG-PVSPPTASANTVNDTIKVGGEVFDGKGTFTAGSELGD 316  
QY 283 GGQSENOKPLFILEDGASLKNVTMGDDGADGIHLGYDAKIDNLHVTNVGEDALTVPNSA 342  
Db 317 GGQAEQKPLPELAQAGATLKNVFGDNDADGVHVRGDAKIDNVHVTNVGEDALTVPNS- 375  
QY 343 GKSHVEITNSSEFHASDKILQNLADTNLSVDNVKAKDFGTFVTRTNGGQGGNDLNLSHI 402  
Db 376 GKPAVETITNSAQASDKIFQLNADANLTIDNFKAKDFGTFVTRTNGGQGGNDLNLNSI 435  
QY 403 SAEDGKFSFKVSDSEGLNVNTSDISLGDVENHYKVPMSANLKV 445  
Db 436 DAENGKFSFKVSDSEGLNVKNINLTNNVNNHYKVPDSANLQV 478  
  
RESULT 7  
Q6D5C8 PRELIMINARY; PRT; 479 AA.  
ID Q6D5C8  
AC Q6D5C8;  
DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Type III effector protein.  
GN Name=hrpW; ORFNames=ECA2112;

OS Pectobacterium atrosepticum.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_TaxID=29471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SCRI1039;  
RA Holvea M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,  
RA Birch P.R.J., Toth I.K.;  
RA Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AV496066; AAS20352.1;  
DR InterPro: IPR011050; Pectin\_lyase\_like.  
DR InterPro: IPR004898; Pect\_lyase.  
DR Pfam: PF03211; Pectate\_lyase; 1.  
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;  
  
Query Match 43.5%; Score 1005; DB 2; Length 479;  
Best Local Similarity 49.9%; Pred. No. 1.1e-43;  
Matches 231; Conservative 60; Mismatches 102; Indels 70; Gaps 15;  
  
QY 8 NNTSSPGLFQSGGDNGLGHNANSALGQOPIDRQITQMAQLLAELLSL----- 59  
Db 61 NTPSSTDG--SSQQAQI-GNGSSALGQN-----GSAADMLKLEALIPSKNGQEA 110  
QY 60 -SQSGNAATGAGN---DQTTGVNAGLNGRKGTTAGTTTQSDSQNMLSEMNGNLDQA 115  
Db 111 GNPLSSGSSGAAGNCGASPLTSGSGAGVGGAQ-----NPEDLSRLSQDSAGSALNNA 165  
QY 116 ITPDQGGGGIGDNPLKMLKLIARMQGSQDFQPGTGNNSASGTSSSGSGSPFNDL 175  
Db 166 INPTADGGGQLSGNDLLKALLELIGNLMDSKQGEFGQP-----QSGSQSGGSP--- 215  
QY 176 SGKAPSGNSPNSGVSVSTFSPSTPTSPSLDFPSSPTKAAG-----STPVTDPD 230  
Db 216 STG-APQASSGGGGSPAA-----PSAPSSVGGNGGAASAPLTAAPT 257  
QY 231 PV--GSAG-----IGAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGTFTAGSELGD 282  
Db 258 GVDGSSAASPASTAGAG-PVSPPTASANTVNDTIKVGGEVFDGKGTFTAGSELGD 316  
QY 283 GGQSENOKPLFILEDGASLKNVTMGDDGADGIHLGYDAKIDNLHVTNVGEDALTVPNSA 342  
Db 317 GGQAEQKPLPELAQAGATLKNVFGDNDADGVHVRGDAKIDNVHVTNVGEDALTVPNS- 375  
QY 343 GKSHVEITNSSEFHASDKILQNLADTNLSVDNVKAKDFGTFVTRTNGGQGGNDLNLSHI 402  
Db 376 GKPAVETITNSAQASDKIFQLNADANLTIDNFKAKDFGTFVTRTNGGQGGNDLNLNSI 435  
QY 403 SAEDGKFSFKVSDSEGLNVNTSDISLGDVENHYKVPMSANLKV 445  
Db 436 DAENGKFSFKVSDSEGLNVKNINLTNNVNNHYKVPDSANLQV 478  
  
RESULT 7  
Q6D5C8 PRELIMINARY; PRT; 479 AA.  
ID Q6D5C8  
AC Q6D5C8;  
DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Type III effector protein.  
GN Name=hrpW; ORFNames=ECA2112;



Db 362 CTNGCKQFNDMSIELNGIDASHGKALVKYSDSEDLKATGDIAMTDVKHAY 412

## RESULT 11

O87327 PRELIMINARY; PRT; 441 AA.  
 AC O87327;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hrp type III secreted protein.  
 GN Name=hrpW;  
 OS Pseudomonas syringae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=317;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B728a;  
 RX MEDLINE=98422476; PubMed=9748456;  
 RA Charkowski A.O., Alfano J.R., Preston G., Yuan J., He S.Y.,  
 RA Collmer A.;  
 RT "The Pseudomonas syringae pv. tomato HrpW protein has domains similar  
 RT to harpins and pectate lyases and can elicit the plant hypersensitive  
 RT response and bind to pectate.";  
 RL J. Bacteriol. 180:5211-5217(1998).  
 DR EMBL: AR037983; AAC62530.1; -.  
 DR HSSP; Q9RHW0; IE66.  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0030570; F:pectate lyase activity; IEA.  
 DR InterPro; IPR011050; Pectin lyas\_like.  
 DR InterPro; IPR004898; Pect lyase.  
 DR Pfam; PF03211; Pectate lyase; 1.  
 DR PIR; P03211; Pectate lyase; 1.  
 SQ SEQUENCE 441 AA; 44714 MW; 7B3B3146E51897BB CRC64;

Query Match 24.6%; Score 567.5; DB 2; Length 441;  
 Best Local Similarity 36.4%; Pred. No. 2.4e-21;  
 Matches 156; Conservative 46; Mismatches 133; Indels 93; Gaps 14;  
 QY 36 QQPIDQTI-----EQMAQLLAELKSLLS---PQSGNAATGAGNDQTTGVNAGLNGR 88  
 DB 67 QNTDSSAATDPQSNVVKLLSALVTSLLQMLMNLKKQDTGQDSNEWQPPFQNEGGLG-- 124  
 QY 89 KGTAGTTTPOSQSONLSEMGNNGLDQAITPDGQGGQIGDNPLLKAMLKLIARMDGQSD 148  
 DB 125 -----TPSA-----EGSDGGT-----QEA 138  
 QY 149 QFQPGTGNNSASGTSSTSGGSPFNDLSGKAPSGNSPS-----GNVSPVST-----FSPPS 200  
 DB 139 SGDEGGGTTAATGDDGGGTSTTEGDDG---GTSPTAEGDGGGYSVSTGADGSGAFS 194  
 QY 201 TPTSPSPDLPSPSPKACAGGSTPVTDHPDP--VGSAGIGAGNSVAFTSAGANQ-----TV 254  
 DB 195 T-----EDGTGGGGSGDGVTPQVTPQANPGRNNGTGVSDTTGSLSEQSEVVV 243  
 QY 255 LHDITIVKAGQVFDGKQGTFTAGSELGDDGQSQSENKQPLFILEDGASLKNVTMGDDGADGI 314  
 DB 244 VKDTIKVGAQGVFDHGATFTADKSMGTGQDEHQKPLPELAGVAILKNVNLGENEADGI 303  
 QY 315 HLYGDAK-----IDNLHVTNVEDAITVKPNAGKSHVEITNSFEHASDKILOINAD 368  
 DB 304 HV--NAKNSQVITDINVAQNVGEMITVKEGGAKVTNLNITNSANGADDKVIQLNAD 361  
 QY 369 TNLSDNVNKAQDFGTFVRTNGGQO--GNWDLNLGSHISAEDGKFSFKVKSDEGLNVNTSDIS 427  
 DB 362 THLVKDGFKATDGTGLVRTNGGQKQFDDMSVELNGVDATHGKFALVKSDSDDLKATGDI 421  
 QY 428 LGDVENHY 435  
 DB 422 MTDVKHAY 429

## RESULT 12

O87264 PRELIMINARY; PRT; 424 AA.  
 AC O87264;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE HrpW.  
 GN Name=hrpW;  
 OS Pseudomonas syringae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=317;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DC3000;  
 RX MEDLINE=98422476; PubMed=9748456;  
 RA Charkowski A.O., Alfano J.R., Preston G., Yuan J., He S.Y.,  
 RA Collmer A.;  
 RT "The Pseudomonas syringae pv. tomato HrpW protein has domains similar  
 RT to harpins and pectate lyases and can elicit the plant hypersensitive  
 RT response and bind to pectate.";  
 RL J. Bacteriol. 180:5211-5217(1998).  
 DR EMBL; AF005221; AAC62526.1; -.  
 DR HSSP; Q9RHW0; IE66.  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0030570; F:pectate lyase activity; IEA.  
 DR InterPro; IPR011050; Pectin lyas\_like.  
 DR InterPro; IPR004898; Pect lyase.  
 DR Pfam; PF03211; Pectate lyase; 1.  
 DR PIR; P03211; Pectate lyase; 1.  
 SQ SEQUENCE 424 AA; 42910 MW; 7AF4ED059BE79D2E CRC64;  
 Query Match 24.2%; Score 559; DB 2; Length 424;  
 Best Local Similarity 36.1%; Pred. No. 6.1e-21;  
 Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
 QY 37 QPIDQTIQEQMAQLLAELKSL---LSPQSGNAATGAGNDQTTGVNAGLNGRKGATAG 93  
 DB 72 KPNDSSQ--NTAKLISALIMSLLOMLTNSNKKQDTNQOPDSQAFQNNGLG----- 122  
 QY 94 TTPQSDSQNMLSEMGNNGLDQAITPDGQGGQIGDNPLLKAMLKLIARMDGQSDQFQOP 153  
 DB 123 -TPSADS-----GGGG-----TPDATGGGG--GDTG-----SATGGG 151  
 QY 154 GTGNNSASGTSSTSGGSPFNDLSGKAPSGNSPSGNYSVPSTFTSPTSPDPPS 213  
 DB 152 GGDITPTATGGGGGGGGTPTATGGG---SGGTPTATGGEGGVTPQITPOL-----A 200  
 QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGT 273  
 DB 201 NPNRTSG-----TGVSVDTAGS-----TEQAGKINNVKDTIKVGAQVFDHGAT 245  
 QY 274 FTAGSELGDDGQSQSENKQPLFILEDGASLKNVTMGDDGADGIHLYG-----DAKIDNLHVTN 329  
 DB 246 FTADKSMGNGDQENKQPFELAEAGATLKNVNLGENEVDGIHVAKNAQAEVITDINVAQN 305  
 QY 330 VGEDAITVKPNAGKSHVEITNSFEHASDKILOINADTNLSVDNVKAKDFGTFVRTNG 389  
 DB 306 VGEDLIITVKEGGAAVTNLNITNSAKGADDKVVQQLNANTHLKIDNFKADDFGTWVRTNG 365  
 QY 390 GQO--GNWDLNLGSHISAEDGKFSFKVKSDEGLNVNTSDISLGDVENHY 435  
 DB 366 GKQFDDMSIELNGIEANHGKFALVKSDSDDLKATGNIAMTDVKHAY 412

## RESULT 13

O7C4I5 PRELIMINARY; PRT; 424 AA.  
 AC O7C4I5; Q7BM42;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Type III helper protein HrpW(Pto) (HrpW).

GN Name=hrpW (Pto); Synonyms=hrpW; OrderedLocusNames=PSPTOL1373;  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;  
RA Buell C.R., Joardar V., Lindeberg M., Selenogut J., Paulsen I.T.,  
RA Gwin M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,  
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,  
RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,  
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,  
RA Kutterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,  
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,  
RA Delaney C.L., White O., Fraser C.M., Collmer A.;  
RA Bender C.L., White O., Fraser C.M., Collmer A.;  
RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
RT Pseudomonas syringae pv. tomato DC3000.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RX MEDLINE=98422476; PubMed=9748456;  
RA Charkoweki A.O., Alfano J.R., Preston G., Yuan J., He S.Y.,  
RA Collmer A.;  
RT "The Pseudomonas syringae pv. tomato HrpW protein has domains similar  
RT to harpins and pectate lyases and can elicit the plant hypersensitive  
RT response and bind to pectate.";  
RL J. Bacteriol. 180:5211-5217(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RX MEDLINE=20243785; PubMed=10781092;  
RA Alfano J.R., Charkowski A.O., Deng W.L., Badel J.L.,  
RA Petnicki-Ocwieja T., van Dijk K., Collmer A.;  
RT "The Pseudomonas syringae Hrp pathogenicity island has a tripartite  
RT mosaic structure composed of a cluster of type III secretion genes  
RT bounded by exchangeable effector and conserved effector loci that  
RT contribute to parasitic fitness and pathogenicity in plants.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4856-4861(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RX MEDLINE=22467693; PubMed=12580281;  
RA Fouts D.E., Badel J.L., Ramos A.R., Rapp R.A., Collmer A.;  
RT "A Pseudomonas syringae pv. tomato DC3000 Hrp (Type III secretion)  
RT deletion mutant expressing the Hrp system of bean pathogen P. syringae  
RT pv. syringae 61 retains normal host specificity for tomato.";  
RL Mol. Plant Microbe Interact. 16:43-52(2003).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RX MEDLINE=22467693; PubMed=12580281;  
RA Fouts D.E., Badel J.L., Ramos A.R., Rapp R.A., Collmer A.;  
RT "A Pseudomonas syringae pv. tomato DC3000 Hrp (Type III secretion)  
RT deletion mutant expressing the Hrp system of bean pathogen P. syringae  
RT pv. syringae 61 retains normal host specificity for tomato.";  
RL Mol. Plant Microbe Interact. 16:43-52(2003).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RX MEDLINE=22467693; PubMed=12580281;  
RA Ramos A.R., Rehm A.H., Collmer A.R.;  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RA Ramos A.R., Rehm A.H., Collmer A.R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF232004; AAF71503.1; -;  
DR EMBL; AF232004; AAF71503.1; -;  
DR InterPro; IPR011050; Pectin\_lyase\_like.  
DR InterPro; IPR004898; Pect\_lyase.  
DR Pfam; PF03211; Pectate\_lyase; 1.  
KW Complete proteome.  
SQ SEQUENCE 424 AA; 42910 MW; 7AF4ED059BE79D2E CRC64;  
Query Match 24.2%; Score 559; DB 2; Length 424;  
Best Local Similarity 36.1%; Pred. No. 6.1e-21;  
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;  
Query Match 20.5%; Score 472.5; DB 2; Length 386;  
Best Local Similarity 32.1%; Pred. No. 1.4e-16;  
Matches 141; Conservative 52; Mismatches 121; Indels 125; Gaps 17;  
Query 16 LPFSGDNGLG-GHNANSALGQPDROT-----TEQAQLLAELKSLSPSGNA 66  
Db 47 LFNSAPKRVSPGPDNSV--QSPQDANNANPENSLKLFSAIVSLQLQMLN-LNKQ 103  
Query 67 ATGAGNDQTTGVGNAGLNGKGTAGTTPQSDSQNMLSEMGNGLDQAITPDGQGGQ 126  
Db 104 DTDQDSSEWQDPFQNGGLG-----TPSDS-----GDSGDSGLAATGDGG--- 145  
Query 127 GNPPLKMLKLIARMMDGQSDQFQPGTGNNSASSGTSSSGSPFNDLSGGKAPSGNSP 186

```

Db 146 GDTT-----TATGNN-----TPSVEGSSND--DGVTPQLANP 176
QY 187 SGNYSPVSTFSPPTSTPTSPSLDPSSPTKAAGGTPVTDHPDPVGSAGIGAGNSVAPT 246
Db 177 N-----HTSGTGPVSD-----187
QY 247 SAGANQ-----TVLHDTITVKAQVDFGKQFTTAGSELGCGQSENOKPLFILEDGASL 301
Db 188 TAGANDQAGKVIWVKTIVKAANTVYDAHGAFTADKSMGNGDQSENOKPLPELAGATL 247
QY 302 KNVTMGGDGDAGIHLYG-DAK---IDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSPEH 357
Db 248 KNAVLENEVDGIHVKAKDAQVQITDNVHAENVGDMITVKGEGAKVTNLQIKNSSAEN 307
QY 358 ASDKILQNLNADTNLSVDNVKADFGFTVRTNGQQ--GNWDLNLSHISAEDGKFSFVKSDS 416
Db 308 ADDKVFQNLNANVTHLNTGPEANNFTGLVRTNGEKQFDDNMNKLGDVANDHGFALVKSDS 367
QY 417 EGLNVNTSISLGDVENHY 435
Db 368 E-----DLQLGDQOHRH 379

RESULT 15
Q8KKZ2
ID Q8KKZ2 PRELIMINARY; PRT; 307 AA.
AC Q8KKZ2;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein hrpW.
GN Name=hrpW;
OS Rhizobium etli.
OG Plasmid symbiotic plasmid p42d.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CFN42;
RX MEDLINE=91193195; PubMed=2013564;
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
RT "Structural complexity of the symbiotic plasmid of Rhizobium
leguminosarum bv. phaseoli.";
RL J. Bacteriol. 173:2411-2419(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CFN42;
RX MEDLINE=97419521; PubMed=9274036;
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
RA Cevallos M.A., Davila G.;
RT "Sequence, localization and characteristics of the replicator region
of the symbiotic plasmid of Rhizobium etli.";
RL Microbiology 143:2825-2831(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=CFN42;
RX MEDLINE=22309397; PubMed=12421308;
RA Quintero V., Cevallos M.A., Davila G.;
RT "A site-specific recombinase (RinQ) is required to exert
incompatibility towards the symbiotic plasmid of Rhizobium etli.";
RL Mol. Microbiol. 46:1023-1032(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=CFN42;
RA Ramirez M.A., Bustos P., Girard L., Rodriguez O., Cevallos M.A.,
RA Davila G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=CFN42;
RA Quintero V., Bustos P., Davila G.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

```

```

RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Gonzalez V., Bustos P., Medrano-Soto A., Ramirez-Romero M.A.,
RA Romero D., Salgado H., Hernandez-Gonzalez I., Hernandez-Cellis J.C.,
RA Quintero V., Girard L.L., Rodriguez O., Flores M., Cevallos M.A.,
RA Collado-Vides J., Davila G.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Quintero V., Bustos P., Davila G.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80928; AAM54988.1; -.
DR HSP; Q9RHW0; IE56.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; E:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate lyase; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 307 AA; 32912 MW; ED8844FA3A3A5E98 CRC64;

Query Match 18.3%; Score 423.5; DB 2; Length 307;
Best Local Similarity 37.7%; Fred.No. 3.6e-14;
Matches 121; Conservative 48; Mismatches 83; Indels 69; Gaps 14;

QY 144 DQSDQFGQPGTGNNSSASGTS---SSGSGSPNDLSGGKAPSGNSPNSVSTFSPPS 200
Db 22 DLESCLG-----GSNRADTDTTITRKVSPEDF-----SGNPPT-----LTSVVPNS 66
QY 201 TPTS-----PTGPLDF-----PSSPTKAAGGTPVTDHPDPVGSAGIGAGN 241
Db 67 RETSENGMDSDPKDLRLKHINWQSDSKVDPSEKQATTLQT-TTEKPD-----114
QY 242 SVAFTSAGANQTVLHDTITVKAQVDFGKGTFTTAGSELGCGQSENOKPLFILEDGASL 301
Db 115 ----LSKSGSVIWNNEPIVWDGG-VFDGKGATYTTASSKLGGGQSETQSPFLIKNGATL 169
QY 302 KNVTMGGDGDAGIHLYGDAKIDNLHVTNVGDEDAITVKPNSAGKSHVEITNSSFHASDK 361
Db 170 KNVDLGENGADGIHYVDGATLENVWQNVGEDALTVK--SAG---DITIGSAGKATDK 224
QY 362 ILQLNADTNLSVDNVKADPGFTVRTNGGQQGNWDLNLSHISAEDGKFS-----FVKSDSE 417
Db 225 IPQINADTRFYLKDFVADGFTLVRTNGGKQIDAD-----VTIDGGAFSHGSNVFRDSS 279
QY 418 GLNVN-TSDISLGDVENHYKV 437
Db 280 LASVTFLSDITLDDVKNWTRV 300

Search completed: January 25, 2005, 16:09:21
Job time : 200 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 19:24:24 ; Search time 47.5096 Seconds  
(without alignments)  
203.868 Million cell updates/sec

Title: DESAI-879\_31\_57

Perfect score: 128

Sequence: 1 NSALGQPIDROTIFQMAQLLAELIKS 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	447	2	Aaw96260 Hypersens
2	128	100.0	447	3	Aay71094 Erwinia a
3	128	100.0	447	3	Aay84855 A hyperse
4	128	100.0	447	5	Aao22548 Hypersens
5	128	100.0	447	5	Aae18296 Erwinia a
6	128	100.0	447	5	Aae16448 E. amylov
7	128	100.0	447	5	Abb9226 Erwinia a
8	52	40.6	313	7	Ab073754 Pseudomon
9	52	40.6	457	7	Ab073936 Pseudomon
10	52	40.6	710	8	Adn73161 Thale cre
11	51	39.8	660	7	Adf29258 Bacterial
12	51	39.8	660	7	Adg32467 Pasteurel
13	51	39.8	1312	2	Aaw22775 Human RAD
14	51	39.8	1312	2	Aaw71295 Human hom
15	51	39.8	1318	5	Abb77985 Amino aci
16	51	39.8	1318	7	Adj68860 Human hea
17	51	39.8	1318	8	Adj66509 RAD50 hom
18	50.5	39.5	266	5	Abg70071 Human pre
19	50.5	39.5	352	5	Aaw22307 Lipase mo
20	50.5	39.5	596	5	Abg61825 Prostate
21	50.5	39.5	596	7	Adb75398 Prostate
22	50	39.1	430	5	Abb47515 Listeria
23	49	38.3	264	7	Ab071596 Pseudomon
24	49	38.3	697	6	Ab407012 Oryza sat
25	49	38.3	3249	6	Abu39648 Protein e

26	48.5	37.9	237	7	ABO70264	Ab070264 Pseudomon
27	48.5	37.9	777	7	ADC03511	Adc03511 Pseudomon
28	48.5	37.9	891	7	ABO76849	Ab076849 Pseudomon
29	48	37.5	223	4	ABB60921	Abb60921 Drosophil
30	48	37.5	320	8	ADH13660	Adh13660 Human ENZ
31	48	37.5	422	6	ABP59216	Abp59216 Human dru
32	48	37.5	492	5	ABG69596	Abg69596 Human NOV
33	48	37.5	492	7	ADJ83019	Adj83019 Human NOV
34	48	37.5	495	6	ABU21525	Abu21525 Protein e
35	48	37.5	630	6	ABU21525	Abu21525 Protein e
36	48	37.5	1419	5	AAO20572	Aao20572 Protein E
37	48	37.5	1419	6	ABB99408	Abb99408 Amino aci
38	48	37.5	1419	8	ADN48494	Adn48494 Human smg
39	48	37.5	1449	7	ADJ70004	Adj70004 Human hea
40	48	37.5	2246	6	ABU41616	Abu41616 Protein e
41	47	36.7	112	7	ABO63216	Ab063216 Klebsiell
42	47	36.7	183	8	ADN47050	Adn47050 Thermococ
43	47	36.7	301	8	ADN97661	Adn97661 S ambofac
44	47	36.7	320	1	ADP80242	Adp80242 Sequence
45	47	36.7	320	1	AAP80714	Aap80714 Sequence

#### ALIGNMENTS

RESULT 1  
AAW96260  
ID AAW96260 standard; protein; 447 AA.  
XX  
AC AAW96260;  
XX  
DT 14-JUN-1999 (first entry)  
XX  
DE Hypersensitive response eliciting protein (HrpW).  
XX  
KW Hypersensitive response elicitor protein; hypersensitive response; hrpW;  
KW pathogen; infection; crop protection; disease resistance;  
KW pest resistance; transgenic plant; colouration; maturation.  
XX  
OS Erwinia amylovora.  
XX  
PN WO9907208-Al.  
XX  
PD 18-FEB-1999.  
XX  
PF 27-JUL-1998; 98WO-US015547.  
XX  
PR 06-AUG-1997; 97US-0055108P.  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
PI Kim JF, Beer SV;  
XX  
DR WPI: 1999-167126/14.  
DR N-PSDB; AAX09007.  
PT New Erwinia amylovora hypersensitive response eliciting gene and protein  
PT - useful for providing transgenic plants and seeds with enhanced growth,  
PT and insect and disease resistance.  
XX  
PS Claim 1; Page 50-51; 54pp; English.  
XX  
CC The hypersensitive response eliciting protein (hrp) or polypeptide is  
CC produced as part of an active defense by plants against incompatible  
CC pathogen infections. The hypersensitive response is a rapid localised  
CC necrosis. The hrp protein and gene when used in nucleotide constructs are  
CC useful for providing disease resistance to plants, insect control to  
CC plants, and enhancing plant growth (enhancing fruit size and earlier  
CC colouration and maturation), by direct application of the protein to  
CC plants, or by producing transgenic plants or seeds using the hrp gene  
XX  
SQ Sequence 447 AA;

Wed Jan 26 08:08:26 2005

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27  
DB 31 NSALGQQPIDRQTIEQMAQLLAELLKS 57

RESULT 2

AA71094  
ID AAY71094 standard; protein; 447 AA.

AC AAY71094;

DT 08-SEP-2000 (first entry)

XX Erwinia amylovora hypersensitive response elicitor #2.

DE Hypersensitive response elicitor; environmental stress resistance; plant.

XX Erwinia amylovora.

OS WO200028055-A2.

FN 18-MAY-2000.

XX 04-NOV-1999; 99WO-US026039.

PF 05-NOV-1998; 98US-0107243P.

XX (SDEN-) EDEN BIOSCIENCE CORP.

PI Wei Z, Schading RL;

DR WPI; 2000-376566/32.

DR N-PSDB; AAD00669.

XX Application of a hypersensitive response elicitor protein to plants to  
PT impart stress resistance.  
XX Disclosure; Page 10-12; 84pp; English.

XX The patent discloses a method to impart stress resistance to plants by  
CC applying a hypersensitive response elicitor in a non-infectious form to a  
CC plant or seed. The present sequence is a hypersensitive response elicitor  
CC protein from Erwinia amylovora. The protein is heat stable, protease  
CC sensitive and suppressed by inhibitors of plant metabolism. It is used to  
CC impart stress resistance to plants

XX SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 3; Length 447;

Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27  
DB 31 NSALGQQPIDRQTIEQMAQLLAELLKS 57

RESULT 3

AA71094  
ID AAY84855 standard; protein; 447 AA.

AC AAY84855;

DT 08-AUG-2000 (first entry)

XX A hypersensitive response elicitor protein.

DE Hypersensitive response; insect control; disease resistance;  
XX Hypersensitive response elicitor; plant growth; vegetable; crop;  
KW hypersensitive response elicitor; plant growth; vegetable; crop;

KW ornamental plant.  
XX Erwinia amylovora.  
OS WO200020452-A2.  
PN 13-APR-2000.  
PD 05-OCT-1999; 99WO-US023181.  
XX 05-OCT-1998; 98US-0103050P.  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
PA Wei Z, Fan H, Niggemeyer JL;  
PI WPI; 2000-303745/26.  
XX N-PSDB; AAA14939.

XX Hypersensitive response elicitor polypeptides useful for imparting  
PT enhanced growth, disease resistance and insect resistance to plants,  
PT especially vegetables and ornamental flowers.

XX Disclosure; Page 12-13; 100pp; English.

XX The present sequence represents a hypersensitive response elicitor  
CC polypeptide. The specification describes hypersensitive response elicitor  
CC polypeptide fragments, which do not elicit a hypersensitive response.  
CC Instead, the proteins impart disease resistance to plants, enhance plant  
CC growth, and/or control insects. The polypeptide fragments may be used to  
CC these properties to plants. The plants which may be treated in this way  
CC include vegetables, crops and ornamental plants such as alfalfa, rice,  
CC wheat, barley, rye, chickory, lettuce, endive, cabbage, brussels sprout,  
CC potato, bean, pea, chickory, cauliflower, broccoli, carrot, squash, pumpkin,  
CC beet, parsnip, turnip, celery, carrot, squash, pumpkin, zucchini,  
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,  
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,  
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis  
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,  
CC carnation or zinnia  
XX SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27  
DB 31 NSALGQQPIDRQTIEQMAQLLAELLKS 57

RESULT 4

AA022548  
ID AAO22548 standard; protein; 447 AA.

AC AAO22548;

DT 28-OCT-2002 (first entry)

XX Hypersensitive response elicitor protein #2 from Erwinia amylovora.  
DE Desiccation inhibitor; longevity enhancer; desiccation inhibitor;  
XX cutting; ornamental plant; hypersensitive response elicitor protein;  
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;  
KW flower bloom; flower.  
XX Erwinia amylovora.

OS WO200237960-A2.

PN 16-MAY-2002.

XX



XX New hypersensitive response elicitor proteins comprising spaced apart  
PT domains having an acidic portion linked to an alpha-helix, useful for  
PT imparting disease or stress resistance, controlling insects or enhancing  
PT plant growth.  
XX  
PS Disclosure; Page 13-14; 99pp; English.  
XX  
PS The patent discloses hypersensitive response elicitor proteins and  
CC nucleotides encoding such proteins. Hypersensitive response elicitor  
CC proteins comprise an isolated pair or more of spaced apart domains, each  
CC comprising an acidic portion linked to an alpha-helix and capable of  
CC eliciting a hypersensitive response in plants. Sequences of the invention  
CC are used to impart disease resistance to plants, to enhance plant growth,  
CC to control insects and/or to impart stress resistance to plants which  
CC includes resistance to environmental stresses such as climate, air  
CC pollution, chemical and nutritional stress. The method of imparting  
CC disease resistance has the potential for treating previously untreatable  
CC diseases, treating diseases systemically and avoiding the use of  
CC infectious agents or environmentally harmful materials. Hyper- sensitive  
CC response elicitor sequences are used to enhance plant growth which  
CC encompasses greater yield, increased in quantity of seeds produced,  
CC percentage of seeds germinated, plant size and biomass, bigger fruits,  
CC earlier fruit coloration and plant maturation. They are also used for  
CC insect control which encompasses preventing direct insect damage to plant  
CC by feeding injury, interfering with insect larval feeding on the plants,  
CC preventing insects from colonizing host plants and releasing phytotoxins.  
CC Sequences of the invention also prevent subsequent disease damage to  
CC plants resulting from insect infection. The present sequence is Erwinia  
CC amylovora hypersensitive response elicitor protein  
XX  
SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27  
|||||  
DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57

## RESULT 7

ABB09226  
ID ABB09226 standard; protein; 447 AA.

XX ABB09226;

XX 08-JUL-2002 (first entry)

XX Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.

XX Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;  
KW postharvest disease.

XX Erwinia amylovora.

XX WO200180639-A2.

XX 01-NOV-2001.

XX 17-APR-2001; 2001WO-US012468.

XX 19-APR-2000; 2000US-0198359P.

XX (EDEN-) EDEN BIOSCIENCE CORP.

XX Wei Z, Qiu D, Remick D;

XX WPI; 2002-041357/05.

XX N-PSDB; ABL51711.

XX Inhibiting post harvest disease (caused by Penicillium, Botrytis,

PT Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a  
PT fruits or vegetables, using hypersensitive response elicitor proteins or  
PT nucleic acids.

XX Disclosure; Page 11-12; 72pp; English.

XX The present invention describes methods for inhibiting post harvest  
CC disease or desiccation and enhancing the longevity in a fruits or  
CC vegetables, using hypersensitive response elicitor proteins or  
CC polypeptides or nucleic acids (I) derived from pathogens (e.g. Erwinia  
CC amylovora, E. Stewartii, E. chrysanthemi, E. carotovora, Xanthomonas,  
CC Pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter).  
CC (I) has bactericidal activity, and can be used in gene therapy. The  
CC method can be used for inhibiting post harvest disease (caused by  
CC Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and  
CC enhancing the longevity in a fruits or vegetables. The method enables  
CC growers, warehouse packers, shippers and suppliers to process, handle and  
CC store fruit and vegetables with reduced losses caused by post harvest  
CC disease and desiccation, therefore reducing costs to the consumer and  
CC improving quality. The present sequence represents a hypersensitive  
CC response elicitor protein given in the exemplification of the present  
CC invention

XX Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27  
|||||

DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57

## RESULT 8

ABB073754  
ID ABB073754 standard; protein; 313 AA.

XX ABB073754;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #5929.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD07325.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 22500; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,



CC	invention.
XX	
SQ	Sequence 710 AA;
	Query Match 40.6%; Score 52; DB 8; Length 710;
	Best Local Similarity 44.4%; Pred. No. 50;
	Matches 12; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
QY	1 NSALGQQPIDRTIEQMQLLAELLS 27          : :       :
Dd	58 NSAFSWQESDRAAAKRAQTOLAEAKN 84          : :       :
RESULT 11	
ADF29258	ID ADF29258 standard; protein; 660 AA.
XX	AC ADF29258;
XX	DT 12-FEB-2004 (first entry)
DE	Bacterial virulence associated Pasteurella multocida protein PM1459.
KW	antibacterial; vaccine; mutant; Gram negative bacterium; mutation;
KW	attenuated virulence; immunogenic composition; passive immunization;
KW	insertion site; transposon tagged mutagenesis.
XX	Pasteurella multocida.
OS	EPI350796-A1.
PN	08-OCT-2003.
PD	05-APR-2002; 2002EP-00290861.
PF	05-APR-2002; 2002EP-00290861.
PR	(MERI-) MERIAL.
XX	Crooke HR, Shea JE, Feldman RG, Goutebroze SG, Le Gros F;
XX	WPI; 2003-781146/74.
DR	N-PSDB; ADF29257.
XX	New attenuated mutant of a Gram-negative bacteria, useful for the
PT	production of immunogenic or vaccine compositions for the prevention of
PT	bacterial infections, particularly Gram negative bacteria.
XX	Example 4; SEQ ID NO 50; 96pp; English.
PS	The invention relates to a mutant Gram negative bacterium (Pasteurella
CC	multocida) comprising a mutation in a nucleotide sequence which codes for
CC	a polypeptide having an identity which is equal or more than 70%, 75%,
CC	80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% with an amino acid sequence
CC	encoded by any of 29 fully defined sequences of 267-2832 bp, given in the
CC	specification, the mutation resulting in attenuated virulence of the
CC	bacterium. The attenuated mutant is useful for the production of
CC	immunogenic or vaccine compositions for the prevention of bacterial
CC	infections, particularly Gram negative bacteria. The antibody preparation
CC	is useful for the production of a passive immunization composition of a
CC	therapeutic composition against Gram-negative bacteria. This sequence
CC	represents the Pasteurella multocida protein PM0773 which is involved in
CC	bacterial virulence.
XX	Sequence 660 AA;
SQ	Query Match 39.8%; Score 51; DB 7; Length 660;
	Best Local Similarity 55.0%; Pred. No. 65;
	Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy	6 QQPIDRTIEQMQLLAELL 25     : :     : :
Dd	178 QOPHNKKIEQLKKIQEELL 197     : :     : :
RESULT 13	
AAM22775	ID AAM22775 standard; protein; 1312 AA.
XX	AC AAM22775;



CC nucleic acid of interest to a pre-determined site, where the nucleic acid  
CC has homology at or around the pre-determined site, in a eukaryote with a  
CC preference for non-homologous recombination. The method comprises  
CC steering an integration pathway towards homologous recombination. The  
CC method is useful for directing integration of a nucleic acid of interest  
CC to a subtelomeric and/or telomeric region in an eukaryote with a  
CC preference for non-homologous recombination. The nucleic acid of interest  
CC comprises an inactive gene to replace an active gene, or vice versa, is a  
CC portion of a gene delivery vehicle, confers a desired property to the  
CC eukaryotic cell, or encodes a therapeutic proteinaceous substance or a  
CC substance conferring resistance for an antibiotic substance to a cell.  
CC The method is useful for improving gene targeting efficiency. The method  
CC is useful in the replacement of an active gene by an inactive gene, for  
CC e.g. for the inactivation of genes controlling undesired side branches of  
CC metabolic pathways, to increase the quality of bulk products such as  
CC starch, or to increase the production of specific secondary metabolites  
CC or to inhibit formation of unwanted metabolites, and also to inactivate  
CC genes controlling senescence in fruits and flowers or that determine  
CC flower pigments. The method is also useful for replacing an inactive gene  
CC by an active gene. For e.g. the replacement of a defective p53 by an  
CC intact p53. Many tumours acquire a mutation in p53 during their  
CC development which renders it inactive and often correlates with a poor  
CC response to cancer therapy. By replacing the defect p53 by an intact p53,  
CC e.g. through gene therapy, conventional anti cancer therapy have better  
CC changes of succeeding. The method is also useful for therapeutic  
CC proteinaceous substance integration. A tumoricidal gene can be delivered  
CC to a pre-determined site present only in e.g. proliferating cells, or  
CC present only in tumour cells, e.g. to the site where a tumour antigen is  
CC expressed form. ABB77984-86 represent RAD50 homologues. RAD50 is involved  
CC in non-homologous recombination  
XX

SQ Sequence 1318 AA;

Query Match 39.8%; Score 51; DB 5; Length 1318;  
Best Local Similarity 47.4%; Pred. NO. 1.4e+02;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDRTIEOMALLAELLKS 27  
          |||:|||||:|:|:|  
Db 1243 LDRENTESLAHALVEIKS 1261

Search completed: January 25, 2005, 20:30:25  
Job time : 54.5096 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 20:14:24 ; Search time 19.7308 Seconds  
(without alignments)  
90.751 Million call updates/sec

Title: DESAI-879\_31\_57

Perfect score: 128

Sequence: 1 NSALGQPIDROTIEQMAQLLAELLS 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pap:\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pap:\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pap:\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pap:\*  
5: /cgn2\_6/prodata/1/iaa/PCUS COMB.pap:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	3	US-09-120-927-2
2	128	100.0	447	4	US-09-431-614-6
3	52	40.6	313	4	US-09-252-991A-22500
4	52	40.6	457	4	US-09-252-991A-22682
5	51	39.8	1312	2	US-08-592-126-148
6	51	39.8	1312	2	US-08-687-080-51
7	51	39.8	1312	4	US-09-168-595-148
8	49	38.3	173	4	US-09-270-767-40996
9	49	38.3	173	4	US-09-270-767-56212
10	49	38.3	264	4	US-09-252-991A-20342
11	48.5	37.9	237	4	US-09-252-991A-19010
12	48.5	37.9	891	4	US-09-252-991A-25595
13	47	36.7	112	4	US-09-489-039A-9733
14	47	36.7	505	4	US-09-252-991A-23615
15	45	35.2	620	3	US-08-982-785A-10
16	45	35.2	620	4	US-09-629-498-10
17	45	35.2	631	4	US-09-252-991A-20418
18	45	35.2	686	3	US-08-982-785A-8
19	45	35.2	686	4	US-09-629-498-8
20	45	35.2	756	3	US-08-982-785A-2
21	45	35.2	756	4	US-09-629-498-2
22	45	35.2	1084	3	US-09-227-725A-3
23	45	35.2	1225	4	US-09-501-171-4
24	45	35.2	1247	4	US-09-501-171-6
25	44.5	34.8	122	4	US-09-248-796A-28164
26	44	34.4	120	4	US-09-107-532A-4295
27	44	34.4	255	4	US-09-252-991A-31162

28	44	34.4	319	1	US-08-125-746-1	Sequence 1, Appli
29	44	34.4	319	4	US-08-948-276-2	Sequence 2, Appli
30	44	34.4	320	1	US-08-125-746-3	Sequence 3, Appli
31	44	34.4	320	4	US-08-948-276-1	Sequence 1, Appli
32	44	34.4	320	6	5225537-4	Patent No. 5225537
33	44	34.4	325	4	US-09-248-796A-20030	Sequence 20030, A
34	44	34.4	325	4	US-09-324-096A-2	Sequence 2, Appli
35	44	34.4	327	3	US-09-324-096A-6	Sequence 4, Appli
36	44	34.4	327	3	US-09-324-096A-4	Sequence 6, Appli
37	44	34.4	328	4	US-09-248-796A-24205	Sequence 24205, A
38	44	34.4	378	4	US-09-710-279-1428	Sequence 1428, Ap
39	44	34.4	378	3	US-09-134-001C-3248	Sequence 3248, Ap
40	44	34.4	457	4	US-08-924-629C-6	Sequence 6, Appli
41	44	34.4	502	4	US-10-101-464A-945	Sequence 945, App
42	44	34.4	635	4	US-09-252-991A-20298	Sequence 20298, A
43	43	33.6	93	4	US-09-489-039A-7880	Sequence 7880, Ap
44	43	33.6	190	4	US-09-252-991A-29854	Sequence 29854, A
45	43	33.6	315	4	US-09-252-991A-31381	Sequence 31381, A

ALIGNMENTS

RESULT 1  
US-09-120-927-2  
; Sequence 2, Application US/09120927  
; Patent No. 6262018  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jihyun Francis  
; APPLICANT: Beer, Steven V.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Nixon, Haigraive, Devans & Doyle LLP  
; STREET: P.O. Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09120,927  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,108  
; FILING DATE: 06-AUG-1977  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1581  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-120-927-2

Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.5e-12;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NSALGQPIDROTIEQMAQLLAELLS 27

```
Db      31 NSALGQPIDRTIEQMAQLLAELLS 57
|||||
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22500
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22682

Query Match      40.6%; Score 52; DB 4; Length 457;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      4 LGQOPIDRTIEQMAQLLAEL 25
|||||
Db      312 LGQSPLLAALAEELLQLLAEM 333
|||||

RESULT 5
US-08-592-126-148
; Sequence 148, Application US/08592126
; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
US-08-592-126-148

Query Match      39.8%; Score 51; DB 2; Length 1312;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY      9 IDRTIEQMAQLLAELLS 27
|||||

Db      31 NSALGQPIDRTIEQMAQLLAELLS 57
|||||
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22500
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22500

Query Match      40.6%; Score 52; DB 4; Length 313;
Best Local Similarity 55.6%; Pred. No. 3.3;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      7 QPIDRTIEQMAQLLAEL 24
|||||
Db      21 QRLERQVVAQVLAEL 38
|||||

RESULT 4
US-09-252-991A-22682
; Sequence 22682, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```

```

/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/168,595
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/592,126
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sholtz, Charles K.
/ REGISTRATION NUMBER: 38,615
/ REFERENCE/DOCKET NUMBER: 4600-0111
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 148:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1312 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
/ US-09-168-595-148

Query Match 39.8%; Score 51; DB 4; Length 1312;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDRQTTEQMAQLIAELLKS 27
   ||: ||: | | | | | | | |
Db 1237 LDRENIESLAHALVEIKS 1255

RESULT 8
US-09-270-767-40996
; Sequence 40996, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40996
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-40996

Query Match 38.3%; Score 49; DB 4; Length 173;
Best Local Similarity 44.4%; Pred. No. 4.9;
Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 NSALGQPIDRQTTEQMAQLIAELLKS 27
   ||: ||: | | | | | | | |
Db 127 NSMLNQPCSQQLLATQLLYARLLRS 153

RESULT 9
US-09-270-767-56212
; Sequence 56212, Application US/09270767
; Patent No. 6703491

```

```
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56212
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-56212

Query Match      38.3%; Score 49; DB 4; Length 173;
Best Local Similarity 44.4%; Pred. No. 4;9;
Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 NSALNQPCSQQLLATLLYARLLRS 153
Db 127 NSMLNQPCSQQLLATLLYARLLRS 153

RESULT 10
US-09-252-991A-20342
; Sequence 20342, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20342
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20342

Query Match      38.3%; Score 49; DB 4; Length 264;
Best Local Similarity 58.8%; Pred. No. 8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 ALGQQPIDRQTIEQMAQ 19
Db 196 AAGQHPDHRVTEQEAQ 212

RESULT 11
US-09-252-991A-19010
; Sequence 19010, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19010
; LENGTH: 237
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19010

Query Match      37.9%; Score 48.5; DB 4; Length 237;
Best Local Similarity 46.2%; Pred. No. 8.4;
Matches 12; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 2 SALGQQ-PIDRQTIEQMAQLLAELLK 26
Db 26 SALPQEKPLPQAFDDDDPQVLAELVR 51

RESULT 12
US-09-252-991A-25595
; Sequence 25595, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25595
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25595

Query Match      37.9%; Score 48.5; DB 4; Length 891;
Best Local Similarity 32.3%; Pred. No. 40;
Matches 10; Conservative 9; Mismatches 3; Indels 9; Gaps 1;

QY 5 GQQPIDRQTIEQMA-----QLLAELLK 26
Db 652 GKQPLEAKTLEEMVTANHVSADQVLPDLR 682

RESULT 13
US-09-489-039A-9733
; Sequence 9733, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9733
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9733

Query Match      36.7%; Score 47; DB 4; Length 112;
Best Local Similarity 45.0%; Pred. No. 6;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 8 PIDRQTIEQMAQLLAELLKS 27
Db 11 PVDRTLAIAIPKLAELTET 30

RESULT 14
```

```

RESULT 15
US-08-982-785A-10
; Sequence 10, Application US/08982785A
; Patent No. 6258929
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
; APPLICANT: Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
; TITLE OF INVENTION: NUCLEIC ACIDS AND DIAGNOSIS USING THEM
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,785A
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,556
; FILING DATE: 02-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29, 066
; REFERENCE/DOCKET NUMBER: 05311/018001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-982-785A-10

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 20:37:16 ; Search time 65.1635 Seconds  
(without alignments)  
149.698 Million cell updates/sec

Title: DESAI-879\_31\_57

Perfect score: 128

Sequence: 1 NSALGQPIDRTIETQMAQLLAELLS 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	9	US-09-835-684-5
2	128	100.0	447	9	US-09-880-371-5
3	128	100.0	447	9	US-09-879-248-6
4	128	100.0	447	15	US-10-010-390-5
5	128	100.0	447	15	US-10-441-736-6
6	128	100.0	447	17	US-10-847-142-5
7	51	39.8	660	15	US-10-406-686A-50
8	51	39.8	1312	14	US-10-393-602-148
9	51	39.8	1318	16	US-10-408-765A-666
10	50.5	39.5	266	14	US-10-043-487-245
11	50.5	39.5	596	14	US-10-205-823-222
12	49	38.3	628	16	US-10-437-963-174471
13	49	38.3	628	14	US-10-183-687-244
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 50, Appli
					Sequence 148, App
					Sequence 666, App
					Sequence 245, App
					Sequence 222, App
					Sequence 174471,
					Sequence 244, App

14	49	38.3	3249	15	US-10-282-122A-67572	Sequence 67572, A
15	48.5	37.9	211	17	US-10-425-115-328898	Sequence 328898,
16	48	37.5	96	17	US-10-425-115-267186	Sequence 267186,
17	48	37.5	470	14	US-10-369-493-4813	Sequence 4813, Ap
18	48	37.5	478	14	US-10-369-493-7572	Sequence 7572, Ap
19	48	37.5	492	14	US-10-032-189-10	Sequence 10, Appl
20	48	37.5	495	14	US-10-032-189-68	Sequence 68, Appl
21	48	37.5	630	15	US-10-282-122A-49449	Sequence 49449, A
22	48	37.5	1419	16	US-10-473-115-2	Sequence 2, Appli
23	48	37.5	1449	16	US-10-408-765A-1810	Sequence 1810, Ap
24	48	37.5	2246	15	US-10-282-122A-69540	Sequence 69540, A
25	47	36.7	83	15	US-10-424-539-147641	Sequence 147641,
26	46	35.9	31	16	US-10-745-069-52	Sequence 52, Appl
27	46	35.9	31	17	US-10-768-288A-52	Sequence 52, Appl
28	46	35.9	31	17	US-10-780-325A-52	Sequence 52, Appl
29	46	35.9	32	16	US-10-869-649-52	Sequence 51, Appl
30	46	35.9	32	16	US-10-745-069-51	Sequence 51, Appl
31	46	35.9	32	17	US-10-768-288A-51	Sequence 51, Appl
32	46	35.9	32	17	US-10-780-325A-51	Sequence 51, Appl
33	46	35.9	32	17	US-10-869-649-51	Sequence 51, Appl
34	46	35.9	33	16	US-10-745-069-50	Sequence 50, Appl
35	46	35.9	33	17	US-10-768-288A-50	Sequence 50, Appl
36	46	35.9	33	17	US-10-780-325A-50	Sequence 50, Appl
37	46	35.9	33	17	US-10-869-649-50	Sequence 50, Appl
38	46	35.9	34	16	US-10-745-069-49	Sequence 49, Appl
39	46	35.9	34	17	US-10-768-288A-49	Sequence 49, Appl
40	46	35.9	34	17	US-10-780-325A-49	Sequence 49, Appl
41	46	35.9	34	17	US-10-869-649-49	Sequence 49, Appl
42	46	35.9	36	16	US-10-745-069-47	Sequence 47, Appl
43	46	35.9	36	17	US-10-768-288A-47	Sequence 47, Appl
44	46	35.9	36	17	US-10-780-325A-47	Sequence 47, Appl
45	46	35.9	36	17	US-10-869-649-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-09-835-684-5  
; Sequence 5, Application US/09835684  
; Patent No. US20020019337A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/71  
; CURRENT APPLICATION NUMBER: US/09/835,684  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-835-684-5

Query Match 100.0%; Score 128; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPIDRTIETQMAQLLAELLS 27

DB 31 NSALGQPIDRTIETQMAQLLAELLS 57

RESULT 2

US-09-880-371-5  
; Sequence 5, Application US/09880371

```
; Patent No. US20020059658A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
; PLANTS
; FILE REFERENCE: 21829/91
; CURRENT APPLICATION NUMBER: US/09/880,371
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/211,585
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-880-371-5

Query Match      100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLKS 57

RESULT 3
US-09-879-248-6
; Sequence 6, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; FILE REFERENCE: 21829/81
; CURRENT APPLICATION NUMBER: US/09/879,248
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/212,211
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-879-248-6

Query Match      100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLKS 57

RESULT 4
US-10-010-390-5
; Sequence 5, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Leon, Ernesto
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; FROM ORNAMENTAL PLANTS
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/248,169
```

---

```
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-010-390-5

Query Match      100.0%; Score 128; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLKS 57

RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; FILE REFERENCE: 21829/203 (BEC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6

Query Match      100.0%; Score 128; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;

QY 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLKS 57

RESULT 6
US-10-847-142-5
; Sequence 5, Application US/10847142
; Publication No. US20040265442A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; DESICCATION
; FILE REFERENCE: 21829/197
; CURRENT APPLICATION NUMBER: US/10/847,142
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/835,684
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
```



```
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-847-142-5

Query Match      100.0%; Score 128; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGOQPIDRQTIEQMAQLLAELLS 27
Db 31 NSALGOQPIDRQTIEQMAQLLAELLS 57

RESULT 7
US-10-406-686A-50
; Sequence 50, Application US/10406686A
; Publication No. US20040033586A1
; GENERAL INFORMATION:
; APPLICANT: CROOKE, HELEN RACHEL
; APPLICANT: SHEA, JACQUELINE ELIZABETH
; APPLICANT: FELDMAN, ROBERT GRAHAM
; APPLICANT: COUTEBROZE, SYLVAIN GABRIEL
; APPLICANT: LEGRAS, FRANCOIS-XAVIER
; TITLE OF INVENTION: ATTENUATED GRAM NEGATIVE BACTERIA
; FILE REFERENCE: 454313-3171.1
; CURRENT APPLICATION NUMBER: US/10/406,686A
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/370,282
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-406-686A-50

Query Match      39.8%; Score 51; DB 15; Length 660;
Best Local Similarity 55.0%; Pred. No. 83;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAELL 25
Db 178 QQPHNKQIEQLKQQL 197

RESULT 8
US-10-393-602-148
; Sequence 148, Application US/10393602
; Publication No. US20030170714A1
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory Polypeptides
; NUMBER OF SEQUENCES: 151.
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/393,602
; FILING DATE: 19-Mar-2003
; CLASSIFICATION DATA:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
```

```
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Rad50, pro-translation of SEQ ID NO:54
; SEQUENCE DESCRIPTION: SEQ ID NO: 148:
US-10-393-602-148

Query Match      39.8%; Score 51; DB 14; Length 1312;
Best Local Similarity 47.4%; Pred. No. 1.8e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDROTIEQMAQLLAELLS 27
Db 1237 LDRENIESLAHALVEIILS 1255

RESULT 9
US-10-408-765A-666
; Sequence 666, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 666
; LENGTH: 1318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-666

Query Match      39.8%; Score 51; DB 16; Length 1318;
Best Local Similarity 47.4%; Pred. No. 1.9e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDROTIEQMAQLLAELLS 27
Db 1243 LDRENIESLAHALVEIILS 1261

RESULT 10
US-10-043-487-245
; Sequence 245, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypepti
; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: B4778A
```



```
RESULT 14
US-10-282-122A-67572
; Sequence 67572, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67572
; LENGTH: 3249
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (650)..(650)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1081)..(1081)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1088)..(1088)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1114)..(1114)
; OTHER INFORMATION: X=any amino acid
; US-10-282-122A-67572

Query Match 38.3%; Score 49; DB 15; Length 3249;
Best Local Similarity 61.1%; Pred. No. 1e+03;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 10 DRQTIEQMAQLLAELLS 27
```

```
Db 2637 DRQTIEQMARHWRNLQA 2654
|||||:|:|
RESULT 15
US-10-425-115-328898
; Sequence 328898, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 328898
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(211)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_63023C.1.pep
; US-10-425-115-328898

Query Match 37.9%; Score 48.5; DB 17; Length 211;
Best Local Similarity 39.4%; Pred. No. 52;
Matches 13; Conservative 4; Mismatches 3; Indels 13; Gaps 1;

QY 3 ALGOQPIDRTIEQ-----MAQLLA 22
|||||:|:|
Db 156 ALGOQPIDRVGLEKEAATLSRMGLRTLAQVLA 188

Search completed: January 25, 2005, 20:58:03
Job time : 66.1635 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 20:02:44 ; Search time 15.3173 Seconds  
(without alignments)  
169.602 Million cell updates/sec

Title: DESAI-879\_31\_57  
Perfect score: 128  
Sequence: 1 NSALGQPIDRTIQMAQLLAELLS 27

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*

- 1: pir1.\*
- 2: pir2.\*
- 3: pir3.\*
- 4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	T18447	HrpW protein - Erw
2	51	39.8	580	G69656	2-succinyl-6-hydro
3	51	39.8	1312	T30845	probable DNA repai
4	50	39.1	430	AE1803	sugar binding prot
5	50	39.1	430	AF1429	sugar binding prot
6	50	39.1	2706	T28155	variant-specific s
7	49	38.3	257	S76537	hypothetical prote
8	49	38.3	902	AH2999	conserved hypothet
9	49	38.3	903	A96284	hypothetical prote
10	48.5	37.9	213	C82953	hypothetical prote
11	48.5	37.9	627	S76462	hypothetical prote
12	48.5	37.9	762	H83348	probable acylase p
13	48	37.5	350	S76312	hypothetical prote
14	47	36.7	246	D87487	uridylyl kinase (
15	47	36.7	356	S01992	glutenin low molec
16	47	36.7	817	T01866	hypothetical prote
17	47	36.7	1122	T28130	hypothetical prote
18	46	35.9	36	PCBO	pancreatic hormone
19	46	35.9	59	1 FCSH	pancreatic hormone
20	46	35.9	75	G91013	hypothetical prote
21	46	35.9	75	A85858	hypothetical prote
22	46	35.9	75	B64988	hypothetical 8.3 k
23	46	35.9	159	A82720	conserved hypothet
24	46	35.9	159	A97502	hypothetical prote
25	46	35.9	409	S76119	hypothetical prote
26	46	35.9	792	T29187	hypothetical prote
27	45.5	35.5	249	A81001	electron transfer
28	45.5	35.5	249	F82018	electron transfer
29	45.5	35.5	929	S75098	hypothetical prote

30	45	35.2	36	1	A61132	pancreatic hormone
31	45	35.2	36	1	C61132	pancreatic hormone
32	45	35.2	36	1	D61132	pancreatic hormone
33	45	35.2	66	1	PCCT	Mut7/mudix family
34	45	35.2	203	2	C95194	mutator protein (i
35	45	35.2	203	2	H98060	conserved hypothet
36	45	35.2	455	2	B72339	hypothetical prote
37	45	35.2	510	2	AG2985	sugar ABC transpor
38	45	35.2	511	2	H98297	hypothetical prote
39	45	35.2	595	2	G96491	conserved hypothet
40	45	35.2	605	2	A82381	polysaccharide
41	45	35.2	705	2	S70691	conserved hypothet
42	45	35.2	978	2	B89971	neural plakophilin
43	45	35.2	1247	2	T42209	hypothetical prote
44	45	35.2	1818	1	S73852	COP9 signalosome c
45	44.5	34.8	397	2	T52302	

ALIGNMENTS

RESULT 1

T18447  
HrpW protein - Erwinia amylovora  
C:Species: Erwinia amylovora  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18447  
R:Gaudriault, S.  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z18936  
A:Accession: T18447  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-447 <CAU>  
A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:e1293451; PIDN:CAA74  
A:Experimental source: strain CFP1430; specific host Pommoideae  
C:Genetics:  
A>Note: hrpW

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPIDRTIQMAQLLAELLS 27

DB 31 NSALGQPIDRTIQMAQLLAELLS 57

RESULT 2

G69656  
2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 4.1.3.-) [validated]  
N:Alternate names: menCF protein, menD protein  
N:Contains: 2-oxoglutarate decarboxylase (EC 4.1.1.71)  
C:Species: Bacillus subtilis  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Aug-2004  
C:Accession: G69656; A61649; S27509; T46640; T46641  
R:Kunat, S.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, J.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: G69656

A;Residues: 1-1312 <KIM>  
A;Cross-references: UNIPROT:P70388; EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AAC528  
C;Genetics:  
A;Gene: RAD50  
A;Map position: 11  
C;Superfamily: RAD50 protein  
C;Keywords: DNA repair

Query Match 39.8%; Score 51; DB 2; Length 1312;  
Best Local Similarity 47.4%; Pred.No.55;  
Matches 9; Conservative 5; Mismatches 0; Gaps 0;

QY 9 IDRTIQEQMAQLLAELLS 27  
:||: |||: :|| |:||  
DB 1237 LDRENIESLAHALVEILKS 1255

RESULT 4  
AE1803  
sugar binding protein (ABC transporter) homolog lin2972 [imported] - Listeria innocua (t  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
R;Accession: AE1803  
C;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AE1803  
A;Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-430 <GLA>  
A;Cross-references: UNIPROT:O926R9; GB:AL592022; PIDN:CAC98197.1; PID:g16415512; GSPDB:C  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin2972

Query Match 39.1%; Score 50; DB 2; Length 430;  
Best Local Similarity 39.1%; Pred.No.23;  
Matches 9; Conservative 7; Mismatches 0; Gaps 0;

QY 5 GQQPIDRTIQEQMAQLLAELLS 27  
:||||: :|| |:||  
DB 341 GAQPYNKEVVEQAKYENEVIKS 363

RESULT 5  
AF1429  
sugar binding protein (ABC transporter) homolog lmo2839 [imported] - Listeria monocytoc  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
R;Accession: AF1429  
C;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AF1429  
A;Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-430 <GLA>  
A;Cross-references: UNIPROT:QBY3J6; GB:NC\_003210; PIDN:CAD01052.1; PID:g16412339; GSPDB  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo2839

Query Match 39.1%; Score 50; DB 2; Length 430;

RESULT 6  
T28155  
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum) (fragmented)  
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)  
C:Species: Plasmodium falciparum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T28155  
R:Rowe, J.A.; Moulds, J.M.; Newbold, C.I.; Miller, L.H.  
Nature 388, 292-295, 1997  
A:Title: Plasmodium falciparum rosetting is mediated by PfEMP1 and requires complement  
A:Reference number: Z20477; MUID:97373957; PMID:9230440  
A:Accession: T28155  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2706 <ROW>  
A:Cross-references: UNIPROT:O15870; EMBL:Y13402; PIDN:CAA73831.1  
A:Experimental source: strain IT 4/25/5  
C:Genetics:  
A:Introns: 2493/3  
A:Note: R29R+var1

Db 2108 GKTPIDANFEOMGOTAKEFV 2128

A/Cross-References: UNIPROT:Q55741; EMBL:D64002; GB:AB001339; NID:g1001612; PIDN:BAA1038  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Db 177 LG000KRRSIOOWAFLLGNLL 198

```
C:\Access: AH2999  
C:\Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
```





Query Match 36.7%; Score 47; DB 2; Length 356;  
Best Local Similarity 58.8%; Pred. No. 50;  
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 LGQOPIDROTIEQMAQL 20  
| | | | : | : | | |  
Db 301 LGQOPQOQLAHQIAQL 317

Search completed: January 25, 2005, 20:36:58  
Job time : 16.3173 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 19:58:54 ; Search time 85.4135 Seconds  
(without alignments)  
181.881 Million cell updates/sec

Title: DESAI-879\_31\_57  
Perfect score: 128  
Sequence: 1 NSALGQQPIDRQTIEQMAQLLAELLS 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	138	2 Q79AW7	Q79AW7 erwinia amy
2	128	100.0	447	2 O54508	O54508 erwinia amy
3	128	100.0	447	2 Q9LAW2	Q9LAW2 erwinia amy
4	109	85.2	450	2 Q6XDB8	Q6XDB8 erwinia pyr
5	109	85.2	450	2 AAQ17046	AAQ17046 erwinia p
6	109	85.2	450	2 AAS45453	AAS45453 erwinia p
7	61.5	48.0	1214	2 Q8DZ21	Q8DZ21 synchococ
8	53	41.4	416	2 Q6DEU7	Q6DEU7 brachydanio
9	53	41.4	440	2 Q6IQ77	Q6IQ77 brachydanio
10	53	41.4	440	2 AAH71533	AAH71533 brachydan
11	53	41.4	943	2 Q7SBG5	Q7SBG5 neurospora
12	52.5	41.0	452	2 Q8U310	Q8U310 pyrococcus
13	51	39.8	377	2 Q7SY40	Q7SY40 brachydanio
14	51	39.8	413	2 Q7NHF2	Q7NHF2 gloebacter
15	51	39.8	498	2 Q8JHT9	Q8JHT9 brachydanio
16	51	39.8	580	1 MEND BACSU	P23970 b menaquin
17	51	39.8	660	2 Q9CKZ1	Q9CKZ1 pasteurella
18	51	39.8	845	2 Q8BTF5	Q8BTF5 mus musculu
19	51	39.8	911	2 Q8BTF7	Q8BTF7 mus musculu
20	51	39.8	1017	2 Q6VG40	Q6VG40 chimpanzee
21	51	39.8	1017	2 AAR02377	AAR02377 chimpanze
22	51	39.8	1173	2 Q9UP86	Q9UP86 homo sapien
23	51	39.8	1312	2 Q92878	Q92878 homo sapien
24	51	39.8	1312	2 P70388	P70388 mus musculu
25	51	39.8	1312	2 Q9JIL8	Q9JIL8 rattus norv
26	51	39.8	1318	2 Q43254	Q43254 homo sapien
27	50.5	39.5	352	1 LIPO PSEWI	O05938 pseudomonas
28	50.5	39.5	596	1 ENH HUMAN	Q96hc4 homo sapien
29	50.5	39.5	596	2 AAT06739	AAT06739 homo sapi
30	50.5	39.5	1263	2 Q7UM20	Q7UM20 rhodopirell
31	50	39.1	430	2 Q926r9	Q926r9 listeria in

32	50	39.1	430	2 Q8Y3J6	Q8Y3J6 listeria mo
33	50	39.1	430	2 Q71VS2	Q71VS2 listeria mo
34	50	39.1	430	2 AAT05594	AAT05594 listeria
35	50	39.1	544	2 Q97473	Q97473 caenorhabdi
36	50	39.1	559	2 Q882U6	Q882U6 pseudomonas
37	50	39.1	2706	2 O15870	O15870 plasmodium
38	49.5	38.7	477	2 Q6FJC5	Q6FJC5 candida gla
39	49	38.3	257	2 Q55741	Q55741 synchocyst
40	49	38.3	305	2 Q88Y51	Q88Y51 lactobacill
41	49	38.3	358	2 Q824U8	Q824U8 chlamydophi
42	49	38.3	457	2 Q7X5J7	Q7X5J7 leuconostoc
43	49	38.3	560	2 Q804Z3	Q804Z3 fugu rubrip
44	49	38.3	628	2 Q84LF4	Q84LF4 oryza sativ
45	49	38.3	687	2 Q81MA8	Q81MA8 drosophila

#### ALIGNMENTS

##### RESULT 1

Q79AW7 PRELIMINARY; PRT; 138 AA.  
 AC Q79AW7;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE HrpW (Fragment).  
 GN Name=hrpW;  
 OS Erwinia amylovora.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Erwinia.  
 OX NCBI\_TaxID=552;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Er321;  
 RX MEDLINE=98115919; PubMed=9448330;  
 RA Bogdanove A.J., Kim J.F., Wei Z., Kolchinsky P., Charkowski A.O.,  
 RA Conlin A.K., Collmer A., Beer S.V.;  
 RT "Homology and functional similarity of an hrp-linked pathogenicity  
 RT locus, dspBF, of Erwinia amylovora and the avirulence locus avrE of  
 RT Pseudomonas syringae pathovar tomato.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330(1998).  
 DR EMBL; U97504; AAC04849.1; -;  
 FT NON\_TER  
 SQ SEQUENCE 138 AA; 13788 MW; 4FE17177C74B3C6 CRC64;  
 Query Match 100.0%; Score 128; DB 2; Length 138;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

##### RESULT 2

O54508 PRELIMINARY; PRT; 447 AA.  
 AC O54508;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DE HrpW protein.  
 GN Name=hrpW;  
 OS Erwinia amylovora.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Erwinia.  
 OX NCBI\_TaxID=552;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CFBP1430;  
 RX MEDLINE=98086111; PubMed=9426142;  
 RA Gaudriault S., Malandrin L., Paulin J.P., Barny M.A.;

RT "DspA, an essential pathogenicity factor of *Erwinia amylovora* showing  
RT homology with AvrE of *Pseudomonas syringae*, is secreted via the Hrp  
RT secretion pathway in a DspB-dependent way.";  
RN Mol. Microbiol. 26:1057-1069(1997).  
RL [2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CFBP1430;  
RX MEDLINE=98316710; PubMed=9654138;  
RA Gaudriault S., Brisset M.N., Barny M.A.;  
RT "HrpW of *Erwinia amylovora*, a new Hrp-secreted protein.";  
RL FEBS Lett. 428:224-228(1998).  
RN [3]  
RN SEQUENCE FROM N.A.  
RC STRAIN=EA321;  
RX STRAIN=EA321;  
RA Kim J.F., Zumoff C.H., Beer S.V.;  
RT "HrpW, a new harpin of *Erwinia amylovora*, is a member of a family of  
RT pectate lyases.";  
RL Phytopathology 87:0-0(1997).  
RN [4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=EA321;  
RX MEDLINE=98422475; PubMed=9748455;  
RA Kim J.F., Beer S.V.;  
RT "HrpW of *Erwinia amylovora*, a new harpin that contains a domain  
RT homologous to pectate lyases of a distinct class.";  
RL J. Bacteriol. 180:5203-5210(1998).  
DR EMBL; Y13831; CAA74158.1; -;  
DR EMBL; U94513; AAC62314.1; -;  
DR PIR; T18447; T18447.  
DR HSP; Q9RHW0; IE66  
DR GO; GO:0005776; C:extracellular; IEA.  
DR GO; GO:0030570; F:pectate lyase activity; IEA.  
DR InterPro; IPR011050; Pectin\_lyas\_like.  
DR InterPro; IPR004898; Pect\_lyase.  
DR Pfam; PF03211; Pectate lyase; 1.  
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 3.8e-09;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPIDRQTIEQMAQLLAELLS 27  
DB 31 NSALGQPIDRQTIEQMAQLLAELLS 57

RESULT 3  
Q9LAW2 PRELIMINARY; PRT; 447 AA.  
AC Q9LAW2;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Harpin HrpW.  
GN Name=hrpW;  
OS *Erwinia amylovora*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Erwinia.  
OX NCBI\_TaxID=552;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EA246;  
RA Kim J.F., Laby R.J., Beer S.V.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF083620; AAF63402.1; -;  
DR HSP; Q9RHW0; IE66.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0030570; F:pectate lyase activity; IEA.  
DR InterPro; IPR011050; Pectin\_lyas\_like.  
DR InterPro; IPR004898; Pect\_lyase.  
DR Pfam; PF03211; Pectate lyase; 1.  
SQ SEQUENCE 447 AA; 45340 MW; 0BBAEA3871EDC2F6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 3.8e-09;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPIDRQTIEQMAQLLAELLS 27  
DB 31 NSALGQPIDRQTIEQMAQLLAELLS 57

RESULT 4  
Q6XDB8 PRELIMINARY; PRT; 450 AA.  
AC Q6XDB8;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE HrpW.  
GN Name=hrpW;  
OS *Erwinia pyrifoliae*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Erwinia.  
OX NCBI\_TaxID=79967;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WT#3;  
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,  
RA Hur J.H., Lim C.K.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WT#3;  
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Cho J.M., Hwang I.,  
RA Hur J.H., Lim C.K.;  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY237642; AAQ17046.1; -;  
DR EMBL; AY530755; AA545453.1; -;  
DR InterPro; IPR011050; Pectin\_lyas\_like.  
DR InterPro; IPR004898; Pect\_lyase.  
DR Pfam; PF03211; Pectate lyase; 1.  
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD32C3A74 CRC64;

Query Match 85.2%; Score 109; DB 2; Length 450;  
Best Local Similarity 88.5%; Pred. No. 1.6e-06;  
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQPIDRQTIEQMAQLLAELLS 26  
DB 31 NSALGQPIDRQTIEQMAQLLAELLS 56

RESULT 5  
AAQ17046 PRELIMINARY; PRT; 450 AA.  
AC AAQ17046;  
DT 03-MAR-2004 (TrEMBLrel. 27, Created)  
DT 03-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 03-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE HrpW.  
GN HRPW.  
OS *Erwinia pyrifoliae*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Erwinia.  
OX NCBI\_TaxID=79967;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WT#3;  
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,  
RA Hur J.H., Lim C.K.;  
RL "Molecular characterization hrp genes cluster of *Erwinia pyrifoliae*  
RT and expression of hrpG encoding elicitor of the hypersensitive  
RT response.";  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
RN EMBL; AY237642; AAQ17046.1; -;

```

SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;
Query Match 85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQPIDROTIEQMAQLLAELLK 26
|||||:|||||:|||||:|||||:|||||
Db 31 NSAQGHPIPDROTIEQMAQLLGELLK 56

RESULT 6
AAS45453 PRELIMINARY; PRT; 450 AA.
ID AAS45453
AC AAS45453;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE HRPW.
DE HRPW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]_TaxID=79967;
RP SEQUENCE FROM N.A.
RC STRAIN=WT3;
RA Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RT "Identification of hrp genes cluster and characterization of HR
RT elicitor hrpNEP gene in Erwinia pyrifoliae.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY530755; AAS45453.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQPIDROTIEQMAQLLAELLK 26
|||||:|||||:|||||:|||||:|||||
Db 31 NSAQGHPIPDROTIEQMAQLLGELLK 56

RESULT 7
Q8DGZ1 PRELIMINARY; PRT; 1214 AA.
ID Q8DGZ1
AC Q8DGZ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hydatonase / oxoprolinase.
GN OrderedLocusNames=tlx2171;
OS Synchococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AF005376; BAC09723.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002821; Hydatonase A.
DR InterPro; IPR003692; Hydatonase B.
DR InterPro; IPR008040; Hydatonase B.
DR Pfam; PF01668; Hydatonase_A; 1.

DR PFam; PF02538; Hydatonase_B; 1.
DR Pfam; PF05378; Hydatonase_A; 1.
DR Complete proteome.
SQ SEQUENCE 1214 AA; 129918 MW; 929531C290F24447 CRC64;

Query Match 48.0%; Score 61.5; DB 2; Length 1214;
Best Local Similarity 54.2%; Pred. No. 15;
Matches 13; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 5 GQPIDROTIE-QMAQLLAELLK 27
|||||:|||||:|||||:|||||:|||||
Db 395 GQPLDRDSVEQPAQLRQEIYRS 418

RESULT 8
Q6DEJ7 PRELIMINARY; PRT; 416 AA.
ID Q6DEJ7
AC Q6DEJ7;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]_TaxID=7955;
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RX PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC077115; AAH77115.1; -.
KW Hypothetical protein.
SQ SEQUENCE 416 AA; 47726 MW; 008C38F5B606CE42 CRC64;

Query Match 41.4%; Score 53; DB 2; Length 416;
Best Local Similarity 38.5%; Pred. No. 79;
Matches 10; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 NSALGQPIDROTIEQMAQLLAELLK 26
|||||:|||||:|||||:|||||:|||||
Db 57 NFGKLGKSTESRVLESQYLIAMLK 82

RESULT 9
Q6IQ77 PRELIMINARY; PRT; 440 AA.
ID Q6IQ77
AC Q6IQ77;
```

```
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Zgc:86915;
GN Name=zgc:86915;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071533; AAH71533.1; -
DR InterPro; IPR01128; Cytochrome_P450.
DR InterPro; IPR002401; EP4501.
DR Pfam; PF00067; P450.1.
DR PRINTS; PR00463; EP4501.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 440 AA; 50312 MW; DID4740B11EFF183 CRC64;

Query Match 41.4%; Score 53; DB 2; Length 440;
Best Local Similarity 42.3%; Pred. No. 84;
Matches 11; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEOMAQLLAELK 26
Db 81 NFGLGKKSIESRVLEESQYLFALLK 106

RESULT 10
ID AAH71533 PRELIMINARY; PRT; 440 AA.
AC AAH71533
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071533; AAH71533.1; -
DR InterPro; IPR01128; Cytochrome_P450.
DR InterPro; IPR002401; EP4501.
DR Pfam; PF00067; P450.1.
DR PRINTS; PR00463; EP4501.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 440 AA; 50312 MW; DID4740B11EFF183 CRC64;

Query Match 41.4%; Score 53; DB 2; Length 440;
Best Local Similarity 42.3%; Pred. No. 84;
Matches 11; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEOMAQLLAELK 26
Db 81 NFGLGKKSIESRVLEESQYLFALLK 106

RESULT 10
ID AAH71533 PRELIMINARY; PRT; 943 AA.
AC AAH71533
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M., Schulte U.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zeiter A., Greenberg D.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Navlor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kanvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plannan M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071533; AAH71533.1; -
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 440 AA; 50312 MW; DID4740B11EFF183 CRC64;

Query Match 41.4%; Score 53; DB 2; Length 440;
Best Local Similarity 42.3%; Pred. No. 84;
Matches 11; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 NSALGQQPIDRQTIEOMAQLLAELK 26
Db 81 NFGLGKKSIESRVLEESQYLFALLK 106

RESULT 11
ID Q7SBG5 PRELIMINARY; PRT; 943 AA.
AC Q7SBG5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU07591.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M., Schulte U.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zeiter A., Greenberg D.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Navlor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kanvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plannan M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
```



RX MEDLINE=22977040; PubMed=14621292;  
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,  
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,  
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,  
RA Takeuchi C., Yamada M., Tabata S.,  
RT "Complete genome structure of *Gloeobacter violaceus* PCC 7421, a  
RT cyanobacterium that lacks thylakoids.";  
RL DNA Res. 10:137-145(2003).  
RL EMBL; AP06577; BAC90526.1; -;  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001431; Insulinase\_like.  
DR InterPro; IPR007863; Peptidase\_M16\_C.  
DR Pfam; PF00675; Peptidase\_M16; 1.  
DR Pfam; PF05193; Peptidase\_M16\_C; 1.  
KW Complete proteome; Protease\_1.  
KW SEQUENCE 413 AA; 45118 MW; 72036E4EA2930737 CRC64;  
SQ  
Query Match 39.8%; Score 51; DB 2; Length 413;  
Best Local Similarity 52.4%; Pred. No. 1.5e+02;  
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 1 NSALGQPIDRQTIEQMAQLL 21  
DB 339 NKLLGQYALGKQNSQVAQLL 359  
RESULT 15  
ID Q8JHT9 PRELIMINARY; PRT; 498 AA.  
AC Q8JHT9;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cytochrome P450.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OC NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Baily A.C.D., Woodin B.R., Stegeman J.J.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
DR EMBL; AF497969; AAM90905.1; -;  
DR HSSP; P00179; 1DT6.  
DR GO; GO:0004497; F:monooxygenase activity; IEA.  
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR InterPro; IPR002401; EP450I.  
DR InterPro; IPR008069; EP450\_CYP2D.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00463; EP450I.  
DR PRINTS; PR01686; EP450ICYP2D.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
KW SEQUENCE 498 AA; 56947 MW; E55CCD63B1AF0A7 CRC64;  
SQ  
Query Match 39.8%; Score 51; DB 2; Length 498;  
Best Local Similarity 34.6%; Pred. No. 1.8e+02;  
Matches 9; Conservative 9; Mismatches 8; Indels 0; Gaps 0;  
QY 1 NSALGQPIDRQTIEQMAQLLAEILK 26  
DB 140 NFGLGKTVEDRVLEESRYLIAELK 165



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 19:24:24 ; Search time 43.9904 Seconds  
(without alignments)  
203.868 Million cell updates/sec

Title: DESAI-879\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDQGGGQIGDNPLKAMKLIA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	AAW96260	Hypersens
2	128	100.0	447	AAW71094	Erwinia a
3	128	100.0	447	AAW84855	A hyperse
4	128	100.0	447	AAO22548	Hypersens
5	128	100.0	447	AAE18296	Erwinia a
6	128	100.0	447	AAE16448	E. amylov
7	128	100.0	447	ABO92226	Erwinia a
8	57	44.5	206	AAW29770	Malassezi
9	54	42.2	837	AAE12381	N-termina
10	54	42.2	837	AAE14656	Murine ST
11	54	42.2	837	ABU10480	Mouse STA
12	52	40.6	135	AAW56974	Partial p
13	52	40.6	135	AAW11979	Mycobacte
14	52	40.6	207	AAW61476	A. fumiga
15	52	40.6	210	ADMS7310	Modular a
16	52	40.6	347	ADMS7298	Modular a
17	49	38.3	110	ADK34437	Novel hum
18	49	38.3	164	ADJ25641	Voltage-g
19	49	38.3	436	ADN73559	Thale cre
20	49	38.3	613	ABR42419	Human pot
21	48.5	37.9	55	ABP02086	Human ORF
22	48.5	37.9	193	ADN99445	Novel hum
23	48.5	37.9	365	ABG60270	Human ova
24	48.5	37.9	365	ABG61741	Novel ova
25	48.5	37.9	392	ADM19994	Protein e

## ALIGNMENTS

## RESULT 1

AAW96260  
ID AAW96260 standard; protein; 447 AA.

XX AAW96260;

XX 14-JUN-1999 (first entry)

XX Hypersensitive response eliciting protein (hrpW).

XX Hypersensitive response elicitor protein; hypersensitive response; hrpW;  
XX pathogen; infection; crop protection; disease resistance;  
XX pest resistance; transgenic plant; colouration; maturation.

XX Erwinia amylovora.

XX WO9907208-A1.

XX 18-FEB-1999.

XX 27-JUL-1998; 98WO-US015547.

XX 06-AUG-1997; 97US-0055108P.

XX (CORR ) CORNELL RES FOUND INC.

XX Kim JF, Beer SV;

XX WPI: 1999-167126/14.

XX N-PSDB; AAX09007.

XX New Erwinia amylovora hypersensitive response eliciting gene and protein  
XX - useful for providing transgenic plants and seeds with enhanced growth,  
XX and insect and disease resistance.

XX Claim 1; Page 50-51; 54pp; English.

XX The hypersensitive response eliciting protein (hrp) or polypeptide is  
XX produced as part of an active defense by plants against incompatible  
XX pathogen infections. The hypersensitive response is a rapid localised  
XX necrosis. The hrp protein and gene when used in nucleotide constructs are  
XX useful for providing disease resistance to plants, insect control to  
XX plants, and enhancing plant growth (enhancing fruit size and earlier  
XX colouration and maturation), by direct application of the protein to  
XX plants, or by producing transgenic plants or seeds using the hrp gene  
XX Sequence 447 AA;

26	48	37.5	43	7	ADN59761	Adn59761 Peptide-v
27	48	37.5	238	5	ABP73460	Abp73460 Candida a
28	48	37.5	505	4	ABB58292	Abb58292 Drosophil
29	47	36.7	72	7	ADF58708	Adf58708 Human pol
30	47	36.7	451	4	ABG10570	Abg10570 Novel hum
31	47	36.7	737	7	ADF83524	Adf83524 C. reinha
32	47	36.7	737	7	ADF83526	Adf83526 C. reinha
33	47	36.7	872	4	ABB69476	Abb69476 Drosophil
34	47	36.7	1343	4	AAU35444	Aau35444 Haemophil
35	47	36.7	1343	6	ABU30272	Abu30272 Protein e
36	47	36.7	1479	4	ABB70936	Abb70936 Drosophil
37	46.5	36.3	353	8	ADI43241	Adi43241 Plant tra
38	46.5	36.3	353	8	ADO62808	Ado62808 Transcrip
39	46.5	36.3	504	4	ABB66600	Abb66600 Drosophil
40	46	35.9	402	3	AAG06450	Aag06450 Arabidops
41	46	35.9	402	6	ADA33693	Ada33693 Acinetoba
42	46	35.9	403	3	AAG43157	Aag43157 Arabidops
43	46	35.9	408	4	AAB79473	Aab79473 Corynebac
44	46	35.9	419	7	ADB70132	Adb70132 C. neofo
45	46	35.9	434	3	AAG06449	Aag06449 Arabidops

```

Query Match      100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAMLKLI A 25
DB 116 ITPDGGGGQIGDNPPLKAMLKLI A 140

RESULT 2
AAAY71094
ID AAAY71094 standard; protein; 447 AA.
XX
AC AAAY71094;
XX
DT 08-SEP-2000 (first entry)
XX
DE Erwinia amylovora hypersensitive response elicitor #2.
XX
KW Hypersensitive response elicitor; environmental stress resistance; plant.
XX
OS Erwinia amylovora.
XX
PN WO200028055-A2.
XX
PD 18-MAY-2000.
XX
PF 04-NOV-1999; 99WO-US026039.
XX
PR 05-NOV-1998; 98US-0107243P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Schading RL;
XX
DR WPI; 2000-376566/32.
XX
DR N-PSDB; AAD00669.
XX
PT Application of a hypersensitive response elicitor protein to plants to
PT impart stress resistance.
XX
PS Disclosure; Page 10-12; 84pp; English.
XX
CC The patent discloses a method to impart stress resistance to plants by
CC applying a hypersensitive response elicitor in a non-infectious form to a
CC plant or seed. The present sequence is a hypersensitive response elicitor
CC protein from Erwinia amylovora. The protein is heat stable, protease
CC sensitive and suppressed by inhibitors of plant metabolism. It is used to
CC impart stress resistance to plants
XX
SQ Sequence 447 AA;

Query Match      100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAMLKLI A 25
DB 116 ITPDGGGGQIGDNPPLKAMLKLI A 140

RESULT 3
AAAY84855
ID AAAY84855 standard; protein; 447 AA.
XX
AC AAAY84855;
XX
DT 08-AUG-2000 (first entry)
XX
DE A hypersensitive response elicitor protein.
XX
KW Hypersensitive response; insect control; disease resistance;
XX hypersensitive response elicitor; plant growth; vegetable; crop;

```

```

KW ornamental plant.
XX
OS Erwinia amylovora.
XX
PN WO200020452-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US023181.
XX
PR 05-OCT-1998; 98US-0103050P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Fan H, Niggemeyer JL;
XX
DR WPI; 2000-303745/26.
XX
DR N-PSDB; AAA14939.
XX
PT Hypersensitive response elicitor polypeptides useful for imparting
PT enhanced growth, disease resistance and insect resistance to plants,
PT especially vegetables and ornamental flowers.
XX
PS Disclosure; Page 12-13; 100pp; English.
XX
CC The present sequence represents a hypersensitive response elicitor
CC polypeptide. The specification describes hypersensitive response elicitor
CC polypeptide fragments, which do not elicit a hypersensitive response.
CC Instead, the proteins impart disease resistance to plants, enhance plant
CC growth, and/or control insects. The polypeptide fragments may be used to
CC these properties to plants. The plants which may be treated in this way
CC include vegetables, crops and ornamental plants such as alfalfa, rice,
CC wheat, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,
CC carnation or zinnia
XX
SQ Sequence 447 AA;

Query Match      100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAMLKLI A 25
DB 116 ITPDGGGGQIGDNPPLKAMLKLI A 140

RESULT 4
AAO22548
ID AAO22548 standard; protein; 447 AA.
XX
AC AAO22548;
XX
DT 28-OCT-2002 (first entry)
XX
DE Hypersensitive response elicitor protein #2 from Erwinia amylovora.
XX
KW Desiccation inhibitor; longevity enhancer; desiccation inhibitor;
KW cutting; ornamental plant; hypersensitive response elicitor protein;
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;
KW flower bloom; flower.
XX
OS Erwinia amylovora.
XX
PN WO200237960-A2.
XX
PD 16-MAY-2002.
XX

```

PT	Improving effectiveness of transgenic plants by topical application of a hypersensitive response elicitor protein to the transgenic plant or by incorporating into the plant a transgene encoding the protein.
XX	Disclosure; Page 13-14; 86pp; English.
XX	The invention relates to methods of improving the effectiveness of transgenic plants which involves either topical application of a hypersensitive response elicitor (HRE) protein to the transgenic plant or incorporating into the transgenic plant a transgene encoding HRE. HRE sequence is used for improving the effectiveness of transgenic plants by maximising the benefit of transgenic traits associated with a deleterious effect on growth, stress tolerance, disease or insect resistance, enhanced growth, herbicide resistance, male sterility, modified flower colour and biochemically modified plant product in the transgenic plants or overcoming the deleterious effects. The present sequence is Erwinia amylovora HRE protein
XX	Sequence 447 AA;
SQ	Query Match 100.0%; Score 128; DB 5; Length 447; Best Local Similarity 100.0%; Pred. No. 5.9e-11; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 ITPDGGGGGQIGDNPLKAMKLIA 25 
Db	116 ITPDGGGGGQIGDNPLKAMKLIA 140 
RESULT 6	
AAE16448	ID AAE16448 standard; protein; 447 AA.
XX	AAE16448;
AC	
XX	
DT	09-APR-2002 (first entry)
XX	
DE	E. amylovora hypersensitive response elicitor protein #2.
XX	
KW	Hypersensitive response elicitor protein; plant growth; fruit coloration; disease resistance; stress resistance; phytotoxin; insect infection; plant maturation.
KW	
XX	
OS	Erwinia amylovora.
XX	
FH	Key Location/Qualifiers
FT	Domain 5..64 /label= Hypersensitive_response_eliciting_domain_1
FT	Region 5..45 /label= Acidic_unit
FT	Region 45..64 /label= Alpha_helix
FT	Domain 103..146
FT	Region 103..131 /label= Hypersensitive_response_eliciting_domain_2
FT	Region 131..146 /label= Acidic_unit
FT	Region 131..146 /label= Alpha_helix
XX	
PN	WO200198501-A2.
XX	
PD	27-DEC-2001.
XX	
PF	12-JUN-2001; 2001WO-US018920.
XX	
PR	16-JUN-2000; 2000US-0212211P.
XX	
PA	(EDEN-) EDEN BIOSCIENCE CORP.
XX	
PI	Fan H, Wei Z;
XX	
DR	WPI; 2002-122282/16.
DR	N-PSDB: AAD27016.

PT Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a  
PT fruits or vegetables, using hypersensitive response elicitor proteins or  
PT nucleic acids.  
XX  
XX Disclosure; Page 11-12; 72pp; English.  
XX  
XX The present invention describes methods for inhibiting post harvest  
CC disease or desiccation and enhancing the longevity in a fruits or  
CC vegetables, using hypersensitive response elicitor proteins or  
CC polypeptides or nucleic acids (I) derived from pathogens (e.g. Erwinia  
CC amylovora, E. Stewartii, E. chrysanthemi, E. carotovora, Xanthomonas,  
CC pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter).  
CC (I) has bactericidal activity, and can be used in gene therapy. The  
CC method can be used for inhibiting post harvest disease (caused by  
CC Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and  
CC enhancing the longevity in a fruits or vegetables. The method enables  
CC growers, warehouse packers, shippers and suppliers to process, handle and  
CC store fruit and vegetables with reduced losses caused by post harvest  
CC disease and desiccation, therefore reducing costs to the consumer and  
CC improving quality. The present sequence represents a hypersensitive  
CC response elicitor protein given in the exemplification of the present  
CC invention  
XX  
XX Sequence 447 AA;  
SQ  
Query Match 100.0%; Score 128; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGIGDNPPLKAMLKLI A 25  
DB 116 ITPDGGGGGIGDNPPLKAMLKLI A 140  
RESULT 8  
AAW29770  
ID AAW29770 standard; protein; 206 AA.  
XX  
XX AAW29770;  
XX 27-AUG-2003 (revised)  
DT 23-FEB-1998 (first entry)  
XX  
XX Malassezia fungus MP-3 antigenic protein.  
DE Malassezia; fungus; antigenic; human; IGE; immunoglobulin E; antibody;  
XX  
XX Malassezia.  
XX  
XX Malassezia.  
OS  
XX WO9721817-A1.  
FN  
XX 19-JUN-1997.  
PD  
XX 10-DEC-1996; 96WO-JP003602.  
PF  
XX 12-DEC-1995; 95JP-00346627.  
PR  
XX 05-SEP-1996; 96JP-00257612.  
PR  
XX 05-SEP-1996; 96JP-00257613.  
XX  
XX (TAKI ) TAKARA SHUZO CO LTD.  
PA  
XX Takesako K, Okado T, Yagihara T, Kuroda M, Onishi Y, Kato I;  
PI Akiyama K, Yasueda H, Yamaguchi H;  
PI  
XX WPI; 1997-332788/30.  
DR  
XX N-PSDB; AAT85876.  
DR  
XX Antigenic proteins from the fungus Malassezia - bind to IGE antibodies  
PT present in patients with Malassezia allergies, useful for diagnosis,  
PT treatment and prevention of such conditions.  
XX  
XX Claim 36; Page 83-84; 162pp; Japanese.  
PS

XX New hypersensitive response elicitor proteins comprising spaced apart  
PT domains having an acidic portion linked to an alpha-helix, useful for  
PT imparting disease or stress resistance, controlling insects or enhancing  
PT plant growth.  
XX  
XX Disclosure; Page 13-14; 99pp; English.  
XX  
XX The patent discloses hypersensitive response elicitor proteins and  
CC nucleotides encoding such proteins. Hypersensitive response elicitor  
CC proteins comprise an isolated pair or more of spaced apart domains, each  
CC comprising an acidic portion linked to an alpha-helix and capable of  
CC eliciting a hypersensitive response in plants. Sequences of the invention  
CC are used to impart disease resistance to plants, to enhance plant growth,  
CC to control insects and/or to impart stress resistance to plants which  
CC includes resistance to environmental stresses such as climate, air  
CC pollution, chemical and nutritional stress. The method of imparting  
CC disease resistance has the potential for treating previously untreatable  
CC diseases, treating diseases systemically and avoiding the use of  
CC infectious agents or environmentally harmful materials. Hyper- sensitive  
CC response elicitor sequences are used to enhance plant growth which  
CC encompasses greater yield, increased in quantity of seeds produced,  
CC percentage of seeds germinated, plant size and biomass, bigger fruits,  
CC earlier fruit coloration and plant maturation. They are also used for  
CC insect control which encompasses preventing direct insect damage to plant  
CC by feeding injury, interfering with insect larval feeding on the plants,  
CC preventing insects from colonising host plants and releasing phytotoxins.  
CC Sequences of the invention also prevent subsequent disease damage to  
CC plants resulting from insect infection. The present sequence is Erwinia  
CC amylovora hypersensitive response elicitor protein  
XX  
XX Sequence 447 AA;  
SQ  
Query Match 100.0%; Score 128; DB 5; Length 447;  
Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGIGDNPPLKAMLKLI A 25  
DB 116 ITPDGGGGGIGDNPPLKAMLKLI A 140  
RESULT 7  
ABB09226  
ID ABB09226 standard; protein; 447 AA.  
XX  
XX ABB09226;  
XX 08-JUL-2002 (first entry)  
DT  
XX  
XX Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.  
DE  
XX Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;  
KW postharvest disease.  
XX  
XX Erwinia amylovora.  
OS  
XX WO200180639-A2.  
FN  
XX 01-NOV-2001.  
PD  
XX 17-APR-2001; 2001WO-US012468.  
PF  
XX 19-APR-2000; 2000US-0198359P.  
PR  
XX (EDEN-) EDEN BIOSCIENCE CORP.  
FA  
XX Wei Z, Qiu D, Remick D;  
PI WPI; 2002-041357/05.  
DR  
XX N-PSDB; ABL51711.  
DR  
XX Inhibiting post harvest disease (caused by Penicillium, Botrytis,  
PT

XX The present sequence represents a specifically claimed antigenic protein  
 CC isolated from the fungus *Malassezia*. The antigenic protein can bind to  
 CC IgE antibodies present in patients with allergic conditions. Antigenic  
 CC proteins, peptides and nucleic acids from the fungus *Malassezia* can be  
 CC used in the diagnosis, treatment and prevention of allergic conditions  
 CC due to *Malassezia* organisms (such as *M. furfur*, *M. sympodialis* and  
 CC *M. pachydermatis*). (Updated on 27-AUG-2003 to correct OS field.)  
 XX Sequence 206 AA;  
 SQ

Query Match 44.5%; Score 57; DB 2; Length 206;  
 Best Local Similarity 45.0%; Pred. No. 2.1;  
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 ITPDGGGGQIGDNPLKAM 20  
 : | : ||| : ||| : :  
 Db 86 LAPQSEGGGQLNDGPLKQAI 105

RESULT 9  
 AAB12381  
 ID AAB12381 standard; peptide; 837 AA.  
 AC AAB12381;  
 XX  
 DT 08-NOV-2000 (first entry)  
 XX  
 DE N-terminal domain of murine STAT-6 protein.  
 XX  
 KW STAT; signal transducer and activator of transcription; crystal;  
 KW drug design; murine.  
 XX  
 OS Mus sp.  
 XX  
 EH Key Location/Qualifiers  
 FT Region 4..9  
 FT /label= Alpha helix 1  
 FT Region 11..18  
 FT /label= Alpha helix 2  
 FT Region 24..29  
 FT /label= Alpha helix 3  
 FT Region 31..36  
 FT /label= Alpha helix 4  
 FT Region 39..43  
 FT /label= Alpha helix 5  
 FT Region 50..73  
 FT /label= Alpha helix 6  
 FT Region 75..90  
 FT /label= Alpha helix 7  
 FT Region 93..113  
 FT /label= Alpha helix 8  
 XX  
 PN US6087478-A.  
 XX  
 PD 11-JUL-2000.  
 XX  
 PF 23-JAN-1998; 98US-00012710.  
 XX  
 PR 23-JAN-1998; 98US-00012710.  
 XX  
 XX (UYRQ ) UNIV ROCKEFELLER.  
 PA  
 XX Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J;  
 XX WPI; 2000-505108/45.  
 DR  
 XX New crystals of an N-terminal fragment of a signal transducer and  
 FT activator of transcription that effectively diffracts x-rays, useful for  
 FT drug screening and development.  
 PT  
 XX Disclosure; Fig 1; 42pp; English.  
 PS  
 XX

CC The present invention relates to a crystal of an N-terminal fragment of a  
 CC signal transducer and activator of transcription (STAT) protein. The  
 CC crystal effectively diffracts x-rays, allowing the determination of the  
 CC atomic coordinates of the N-terminal domain to a resolution of greater  
 CC than 5.0 Angstroms. The present sequence is the N-terminal domain of the  
 CC murine STAT 6 protein. The N-terminal domain enables STAT dimers to  
 CC interact and bind DNA cooperatively, a mechanism important for gene  
 CC activation. The crystals are useful in drug screening and development by  
 CC selecting a potential drug by performing rational drug design with the 3-  
 CC dimensional structure determined for the crystal  
 XX  
 SQ Sequence 837 AA;  
 Query Match 42.2%; Score 54; DB 3; Length 837;  
 Best Local Similarity 64.3%; Pred. No. 30;  
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 5 GGGGGQIGDNPLLK 18  
 | : ||| : ||| : ||| :  
 Db 802 GEGGSLGSGQLLK 815

RESULT 10  
 AAE14656  
 ID AAE14656 standard; protein; 837 AA.  
 XX  
 AC AAE14656;  
 XX  
 DT 16-JUL-2002 (first entry)  
 XX  
 DE Murine STAT6 protein.  
 XX  
 KW Signal transducer and activator of transcription; STAT6;  
 KW drug development; drug discovery; crystal; inflammation; allergy; asthma;  
 KW leukaemia; anaemia; neutropaenia; thrombocytopaenia; cancer; obesity;  
 KW viral disease; growth retardation; murine.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..125  
 FT /note= "Conserved N-terminal domain of the STAT family"  
 XX  
 PN US6312887-B1.  
 XX  
 PD 06-NOV-2001.  
 XX  
 PF 24-APR-2000; 2000US-00556273.  
 XX  
 PR 23-JAN-1998; 98US-00012710.  
 XX  
 XX (UYRQ ) UNIV ROCKEFELLER.  
 PA  
 XX Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J;  
 XX WPI; 2002-033337/04.  
 DR  
 XX Identifying compounds that bind to signal transducer and activator of  
 PT transcription proteins, useful for the production of new drugs.  
 PT  
 XX Example; Col 63-69; 44pp; English.  
 PS  
 XX The invention relates to methods for detecting compounds that bind to  
 CC signal transducer and activator of transcription (STAT) proteins for the  
 CC discovery and development of new drug compounds based on the structural  
 CC properties of the protein crystal. The methods include: identifying a  
 CC compound that binds to the N-terminal domain of a STAT protein,  
 CC identifying a compound that enhances or diminishes the binding of the  
 CC dimeric STAT proteins to each other and/or their nucleic acid binding  
 CC site; or identifying a compound that enhances or diminishes the ability  
 CC of STAT protein dimers to induce the expression of a gene operably under  
 CC the control of a promoter containing at least two adjacent weak binding  
 CC sites for STAT protein dimers. The methods are used for identifying new

CC drugs. An antagonist of STAT N-terminal dimeric interactions that  
CC inhibits the binding of the STAT dimers to adjacent weak binding sites on  
CC a promoter of a gene, could be useful as drugs in the treatment of  
CC diseases, e.g. inflammation, allergy, asthma and leukaemias. On the other  
CC hand, an agonist of N-terminal dimeric interactions between STAT dimers,  
CC can be used as drugs in the treatment of diseases e.g. anaemia,  
CC neutropaenia, thrombocytopaenia, cancer, obesity, viral diseases and  
CC growth retardation. The present sequence is murine STAT6 protein  
XX  
XX Sequence 837 AA;  
SQ  
Query Match 42.2%; Score 54; DB 5; Length 837;  
Best Local Similarity 64.3%; Pred. No. 30;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 5 GGGGGQIGDNPPLK 18  
Db 802 GGGGSLGSQLK 815  
RESULT 11  
ABU10480  
ID ABU10480 standard; protein; 837 AA.  
XX  
AC ABU10480;  
XX  
DT 06-AUG-2003 (first entry)  
XX  
DE Mouse STAT6 protein.  
XX  
KW Mouse; signal transducer and activator of transcription; drug design;  
KW drug screening; STAT-STAT dimer interaction; STAT6.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Region 4.  
FT /label= alpha\_helix\_1  
FT Region 11.  
FT /label= alpha\_helix\_2  
FT Region 24.  
FT /label= alpha\_helix\_3  
FT Region 31.  
FT /label= alpha\_helix\_4  
FT Region 39.  
FT /label= alpha\_helix\_5  
FT Region 50.  
FT /label= alpha\_helix\_6  
FT /note= "Residues 57, 61, 65, 68 and 71 contribute to  
FT packing of the coiled-coil"  
FT 75.  
FT /label= alpha\_helix\_7  
FT /note= "Residues 77, 81, 84 and 88 contribute to packing  
FT of the coiled-coil"  
FT 93.  
FT /label= alpha\_helix\_8  
XX  
FN US2003003563-A1.  
XX  
XX 02-JAN-2003.  
XX  
XX 19-OCT-2001; 2001US-00045792.  
XX  
XX 23-JAN-1998; 98US-00012710.  
XX 24-APR-2000; 2000US-00556273.  
XX (VINK/) VINKMEIER U.  
XX (MOAR/) MOAREFI I.  
XX (DARN/) DARNELL J E.  
XX (KURI/) KURIYAN J.  
XX Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J;  
XX

DR WPI; 2003-447354/42.  
XX New crystal having an N-terminal domain of a STAT protein performing X-  
PT ray crystallographic studies, useful for screening drugs that enhance or  
PT inhibit STAT-STAT dimer interactions.  
XX  
PS Disclosure; Page 33-35; 46pp; English.  
XX  
CC The invention relates to a crystal of an N-terminal domain of signal  
CC transducer and activator of transcription (STAT) protein, where the  
CC crystal effectively diffracts X-rays for the determination of the atomic  
CC coordinates of the N-terminal domain of the STAT protein to a resolution  
CC of greater than 5.0 Angstrom. The methods and compositions are useful for  
CC the design and screening of drugs that enhance or inhibit STAT-STAT dimer  
CC interactions. The present sequence represents the amino acid sequence of  
CC mouse STAT6 protein  
XX  
XX Sequence 837 AA;  
SQ  
Query Match 42.2%; Score 54; DB 6; Length 837;  
Best Local Similarity 64.3%; Pred. No. 30;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 5 GGGGGQIGDNPPLK 18  
Db 802 GGGGSLGSQLK 815  
RESULT 12  
AAR56974  
ID AAR56974 standard; protein; 135 AA.  
XX  
AC AAR56974;  
XX  
DT 28-FEB-1995 (first entry)  
XX  
DE Partial PMGA 1.3 protein of Mycoplasma gallisepticum.  
XX  
KW PMGA; adhesin gene complex; hemagglutinin; conserved sequences; primers;  
KW probes; amplification; polymerase chain reaction; specific; detection;  
KW PCR; 1.2; 1.3.  
XX  
OS Mycoplasma gallisepticum.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..25  
FT /label= putative signal peptide  
FT Protein 26..135  
FT /label= mature pMGAL.2 protein  
XX  
FN AU9350593-A.  
XX  
XX 26-MAY-1994.  
XX  
XX 10-NOV-1993; 93AU-00050593.  
XX  
XX 10-NOV-1992; 92AU-00005744.  
XX  
XX (UYME ) UNIV MELBOURNE.  
XX  
XX Browning GF, Markham PF, Whithear KG, Walker ID, Glew MD;  
XX WPI; 1994-209061/26.  
XX N-PSDB; AAQ68668.  
XX Recombinant DNA constructs for Mycoplasma gallisepticum - for diagnosis,  
XX treatment and prophylaxis of poultry respiratory disorders.  
XX Example 1; Fig 3; 51pp; English.  
XX  
XX AAR56974 shows the partial amino acid sequence of pMGAL.3 derived from  
CC the partial DNA sequence of a 10 kb insert from a Mycoplasma  
CC gallisepticum genomic DNA library, detected by probes based on the T3 and  
CC



XX 01-JUL-2004 (first entry)  
XX Modular antigen transporter molecule protein SEQ ID NO: 16.  
DE modular antigen transporter molecule; MAT molecule; immunosuppressive;  
XX anti-allergic; antirheumatic; virucide; antibacterial; cytostatic;  
KW translocation module; targeting module; antigen module.  
XX  
XX Aspergillus fumigatus.  
XX  
XX EP1408114-A1.  
XX  
XX 14-APR-2004.  
XX  
XX 11-OCT-2002; 2002EP-00022774.  
XX  
XX 11-OCT-2002; 2002EP-00022774.  
XX (BIOV-) BIOVISION AG.  
XX  
XX Crameri R, Flueckiger S, Lamping N, Daigle I;  
PI  
XX WPI; 2004-307083/29.  
XX  
XX N-PSDB; ADM57309.  
XX  
XX Modular antigen-transporting molecule, useful for treating, preventing  
PT and diagnosing e.g. autoimmune disease, comprises separate translocation,  
PT targeting and antigen modules.  
XX  
XX Disclosure; Page 48-49; 69pp; German.  
XX  
XX The present invention relates to a modular antigen-transporting molecule  
CC (MAT) comprising one each of a translocation module, a targeting module  
CC and an antigen module. MAT molecules and their coding sequences are used  
CC for imaging and as pharmaceutical, vaccinating and diagnostic agents,  
CC for preventing, reducing and/or stimulating the immune response, and for  
CC treatment of autoimmune, allergic, rheumatism, organ rejection, infection  
CC (bacterial, viral or caused by eukaryotic pathogens) and/or malignant  
CC disease. The present sequence is a MAT molecule shown in the  
XX exemplification of the invention.  
XX  
XX Sequence 210 AA;  
SQ  
Query Match 40.6%; Score 52; DB 8; Length 210;  
Best Local Similarity 40.9%; Pred. No. 13;  
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
QY 1 ITPDGGGGGIGDNPFLKMLK 22  
Db 84 LAPEKSGGKIDQAPVLKAAIE 105  
Search completed: January 25, 2005, 20:30:27  
Job time : 45.9904 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 20:14:24 ; Search time 18.2692 Seconds  
(without alignments)  
90.751 Million cell updates/sec

Title: DESAI-879\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDQGGGGQIGDNPFLKAMKLIA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	3	US-09-120-927-2
2	128	100.0	447	4	US-09-431-614-6
3	57	44.5	206	4	US-09-091-097-6
4	54	42.2	837	3	US-09-012-710-12
5	54	42.2	837	3	US-09-556-273-12
6	51	39.8	409	3	US-09-140-465-3
7	48	37.5	252	4	US-09-248-796A-18831
8	48	37.5	256	4	US-09-270-767-46260
9	47	36.7	379	2	US-08-576-626A-80
10	46	35.9	402	4	US-09-328-352-4980
11	46	35.9	502	4	US-09-270-767-45060
12	46	35.9	749	2	US-08-568-459A-6
13	46	35.9	749	2	US-08-487-826B-6
14	46	35.9	749	3	US-09-210-288-6
15	46	35.9	1098	4	US-09-252-991A-20186
16	45	35.2	82	4	US-09-248-796A-22679
17	45	35.2	133	4	US-09-543-681A-6484
18	45	35.2	230	4	US-09-270-767-39647
19	45	35.2	230	4	US-09-270-767-54864
20	45	35.2	365	3	US-09-151-592-2
21	45	35.2	365	3	US-09-254-077A-10
22	45	35.2	365	4	US-09-443-766-13
23	45	35.2	373	3	US-09-254-077A-12
24	45	35.2	427	4	US-09-270-767-42751
25	45	35.2	515	2	US-08-705-660-46
26	45	35.2	515	3	US-08-989-045-45
27	45	35.2	664	4	US-09-538-092-838

28	45	35.2	1403	4	US-09-252-991A-25500	Sequence 25500, A
29	44.5	34.8	1151	4	US-09-328-352-5876	Sequence 5876, Ap
30	44	34.4	335	4	US-09-252-991A-24116	Sequence 24116, A
31	44	34.4	399	4	US-09-489-039A-14312	Sequence 14312, A
32	44	34.4	690	4	US-09-248-796A-19169	Sequence 19169, A
33	44	34.4	816	4	US-09-540-236-3443	Sequence 3443, Ap
34	43.5	34.0	149	3	US-09-227-357-136	Sequence 196, Appl
35	43.5	34.0	415	3	US-09-100-193-2	Sequence 2, Appl
36	43	33.6	359	4	US-09-252-991A-18134	Sequence 18134, A
37	43	33.6	450	4	US-09-252-991A-17345	Sequence 27345, A
38	43	33.6	515	4	US-09-252-991A-32251	Sequence 32251, A
39	43	33.6	619	4	US-09-252-991A-32940	Sequence 32940, A
40	43	33.6	706	4	US-09-270-767-40994	Sequence 40994, A
41	43	33.6	706	4	US-09-270-767-56210	Sequence 56210, A
42	43	33.6	975	4	US-09-328-352-4764	Sequence 4764, Ap
43	42.5	33.2	197	4	US-09-270-767-32061	Sequence 32061, A
44	42.5	33.2	197	4	US-09-270-767-47278	Sequence 47278, A
45	42.5	33.2	262	4	US-09-248-796A-19720	Sequence 19720, A

## ALIGNMENTS

RESULT 1  
US-09-120-927-2  
; Sequence 2, Application US/09120927  
; Patent No. 6262018  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jihyun Francis  
; APPLICANT: Beer, Steven V  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hartgrave, Devans & Doyle LLP  
; STREET: P O Box 1051, Clinton Square  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,927  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/055,108  
; FILING DATE: 06-AUG-1977  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1581  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-120-927-2

Query Match 100.0%; Score 128; DB 3; Length 447;  
Best Local Similarity 100.0%; Pred. No. 2.5e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ITPDQGGGGQIGDNPFLKAMKLIA 25

	TELEPHONE: 703-205-8000	
	TELEFAX: 703-205-8050	
	INFORMATION FOR SEQ ID NO: 6:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 206 amino acids	
	TYPE: amino acid	
	TOPOLOGY: linear	
	MOLECULE TYPE: protein	
	US-09-091-097-6	
	Query Match 44.5%; Score 57; DB 4; Length 206;	
	Best Local Similarity 45.0%; Pred. No. 0.6;	
	Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;	
QY	1 ITPDGGGQGIGDNPLLKAM 20	
DB	86 LAPQSEGGQLNDGFLKQAI 105	
	RESULT 4	
	US-09-012-710-12	
	; Sequence 12, Application US/09012710	
	; Patent No. 6087478	
	; GENERAL INFORMATION:	
	; APPLICANT: Vinkemeier, Uwe	
	; APPLICANT: Moarefi, Ismail	
	; APPLICANT: Darnell, Jr., James E.	
	; APPLICANT: Kuriyan, John	
	; TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A	
	; TITLE OF INVENTION: STAT PROTEIN AND METHODS OF USE THEREOF	
	; NUMBER OF SEQUENCES: 13	
	; CORRESPONDENCE ADDRESS:	
	; ADDRESSEE: Klauber & Jackson	
	; STREET: 411 Hackensack Avenue, 4th Floor	
	; CITY: Hackensack	
	; STATE: New Jersey	
	; COUNTRY: USA	
	; ZIP: 07601	
	; COMPUTER READABLE FORM:	
	; MEDIUM TYPE: Floppy disk	
	; COMPUTER: IBM PC compatible	
	; OPERATING SYSTEM: PC-DOS/MS-DOS	
	; SOFTWARE: Patent In Release #1.0, Version #1.30	
	; CURRENT APPLICATION DATA:	
	; APPLICATION NUMBER: US/09/012,710	
	; FILING DATE:	
	; CLASSIFICATION:	
	; ATTORNEY/AGENT INFORMATION:	
	; NAME: Jackson Esq., David A.	
	; REGISTRATION NUMBER: 26,742	
	; REFERENCE/DOCKET NUMBER: 600-1-194	
	; TELECOMMUNICATION INFORMATION:	
	; TELEPHONE: 201-487-5800	
	; TELEFAX: 201-343-1684	
	; TELEX: 133521	
	; INFORMATION FOR SEQ ID NO: 12:	
	; SEQUENCE CHARACTERISTICS:	
	; LENGTH: 837 amino acids	
	; TYPE: amino acid	
	; STRANDEDNESS: single	
	; TOPOLOGY: linear	
	; MOLECULE TYPE: protein	
	; HYPOTHETICAL: NO	
	US-09-012-710-12	
	Query Match 42.2%; Score 54; DB 3; Length 837;	
	Best Local Similarity 64.3%; Pred. No. 8.1;	
	Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
QY	5 GGCGGQIGDNPLLK 18	
DB	802 GEGGGSLGSQPLLK 815	

RESULT 5  
US-09-556-273-12  
; Sequence 12, Application US/09556273  
; Patent No. 6312887  
; GENERAL INFORMATION:  
; APPLICANT: Vinkemeier, Uwe  
; APPLICANT: Moarefi, Ismail  
; APPLICANT: Darnell, Jr., James E.  
; APPLICANT: Kuriyan, John  
; TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A  
; TITLE OF INVENTION: STAT PROTEIN AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/556,273  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: 09/012,710  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-194  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 837 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-09-556-273-12

Query Match 42.2%; Score 54; DB 3; Length 837;  
Best Local Similarity 64.3%; Pred. No. 8.1;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 5 GGGGGQIGDNPILK 18  
DB 802 GGGGSLGSQPLK 815

RESULT 6  
US-09-140-466-3  
; Sequence 3, Application US/09140466  
; Patent No. 6268160  
; GENERAL INFORMATION:  
; APPLICANT: CLOUGH, BARBARA  
; APPLICANT: PREISER, PETER  
; APPLICANT: WILSON, ROBERT  
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE  
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS  
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS  
; FILE REFERENCE: N68637B GCW EJC DP  
; CURRENT APPLICATION NUMBER: US/09/140,466  
; CURRENT FILING DATE: 1998-08-26

; EARLIER APPLICATION NUMBER: US 60/056,246  
; EARLIER FILING DATE: 1997-08-28  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Anacystis nidulans  
US-09-140-466-3

Query Match 39.8%; Score 51; DB 3; Length 409;  
Best Local Similarity 57.1%; Pred. No. 10;  
Matches 12; Conservative 3; Mismatches 4; Indels 2; Gaps 1;  
QY 6 QGG--QGIDNPLKMLKLI 24  
DB 183 QGGASGQGDNPWDKILKLM 203

RESULT 7  
US-09-248-796A-18831  
; Sequence 18831, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 18831  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-18831

Query Match 37.5%; Score 48; DB 4; Length 252;  
Best Local Similarity 43.5%; Pred. No. 17;  
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;  
QY 2 TPDGGGGQIGDNPILKMLKLI 24  
DB 64 TTDGNGGGGIGIVPTLQIVATV 86

RESULT 8  
US-09-270-767-46260  
; Sequence 46260, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 46260  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-46260

Query Match 37.5%; Score 48; DB 4; Length 256;  
Best Local Similarity 47.6%; Pred. No. 18;  
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
QY 4 DQGGGGQIGDNPILKMLKLI 24

```
; SEQ ID NO 4980
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4980

Query Match      35.9%; Score 46; DB 4; Length 402;
Best Local Similarity 39.1%; Pred. No. 59;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPLKMLK 23
Db 147 VAPDYQGLTGLGHPYMAKSQL 169

RESULT 11
US-09-270-767-45060
; Sequence 45060, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45060
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45060

Query Match      35.9%; Score 46; DB 4; Length 502;
Best Local Similarity 56.2%; Pred. No. 75;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 PDGGGGQIGDNPLK 18
Db 405 PDTGGGGAGDRPSIK 420

RESULT 12
US-08-568-459A-6
; Sequence 6, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995

; SEQ ID NO 4980
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4980

Query Match      36.7%; Score 47; DB 2; Length 379;
Best Local Similarity 56.2%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPL 16
Db 189 IISGEGGAVITDNPV 204

RESULT 10
US-09-328-352-4980
; Sequence 4980, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Berton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
```

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
;
US-08-568-459A-6

Query Match 35.9%; Score 46; DB 2; Length 749;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNPLLK 18
Db 432 TNSSGGGNYGDRQISK 448

RESULT 14
US-09-210-288-6
; Sequence 6, Application US/09210288
; Patent No. 6392026
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
;
US-09-210-288-6

Query Match 35.9%; Score 46; DB 3; Length 749;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNPLLK 18
Db 432 TNSSGGGNYGDRQISK 448

RESULT 15
```

US-09-252-991A-20186  
; Sequence 20186, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20186  
; LENGTH: 1098  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20186

Query Match 35.9%; Score 46; DB 4; Length 1098;  
Best Local Similarity 47.6%; Pred. No. 1.8e+02;  
Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 1 ITPDGGGGGQ---IGDNPLL 17  
:|||||:|:|  
Db 594 VQPDGGGGAQRHRHAGEGPHL 614

Search completed: January 25, 2005, 20:38:21  
Job time : 19.2692 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 20:37:16 ; Search time 60.3365 Seconds  
(without alignments)  
149.698 Million cell updates/sec

Title: DESAI-879\_116\_140

Perfect score: 128

Sequence: 1 ITPDQGGGGQIGDNPFLKAMKLIA 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	9	US-09-835-684-5
2	128	100.0	447	9	US-09-880-371-5
3	128	100.0	447	9	US-09-879-248-6
4	128	100.0	447	14	US-10-010-390-5
5	128	100.0	447	15	US-10-441-736-6
6	128	100.0	447	17	US-10-847-142-5
7	57	44.5	206	14	US-10-103-670-10
8	54	42.2	837	14	US-10-045-792-12
9	51	39.8	409	9	US-09-845-335-3
10	50	39.1	78	15	US-10-424-599-214915
11	50	39.1	177	16	US-10-437-963-149121
12	49	38.3	149	17	US-10-425-115-231795
13	49	38.3	483	14	US-10-369-493-3880

14	48.5	37.9	365	9	US-09-908-711-100	Sequence 100, App
15	48	37.5	43	14	US-10-269-806-163	Sequence 163, App
16	48	37.5	111	16	US-10-437-963-124136	Sequence 124136, App
17	48	37.5	238	14	US-10-032-585-7297	Sequence 7297, App
18	47.5	37.1	207	17	US-10-425-115-301902	Sequence 301902, App
19	47	36.7	81	17	US-10-425-115-347563	Sequence 347563, App
20	47	36.7	102	15	US-10-424-599-216781	Sequence 216781, App
21	47	36.7	134	17	US-10-425-115-351040	Sequence 351040, App
22	47	36.7	213	17	US-10-425-115-210287	Sequence 210287, App
23	47	36.7	435	14	US-10-369-493-6514	Sequence 6514, App
24	47	36.7	1343	9	US-09-815-242-11037	Sequence 11037, App
25	47	36.7	1343	15	US-10-282-122A-58196	Sequence 58196, App
26	46.5	36.3	353	15	US-10-374-780A-1704	Sequence 1704, App
27	46.5	36.3	353	16	US-10-437-963-194124	Sequence 194124, App
28	46.5	36.3	354	16	US-10-437-963-195060	Sequence 195060, App
29	46	35.9	408	16	US-10-781-014-462	Sequence 462, App
30	46	35.9	419	15	US-10-320-797-3176	Sequence 3176, App
31	46	35.9	441	9	US-09-738-626-5934	Sequence 5934, App
32	46	35.9	450	17	US-10-425-115-332466	Sequence 332466, App
33	46	35.9	469	15	US-10-425-114-57241	Sequence 57241, App
34	46	35.9	749	13	US-10-153-273-6	Sequence 6, Appl
35	46	35.9	1086	9	US-09-924-154-15	Sequence 15, Appl
36	45.5	35.5	113	15	US-10-424-599-183658	Sequence 183658, App
37	45.5	35.5	161	14	US-10-369-493-5035	Sequence 5035, App
38	45.5	35.5	880	16	US-10-437-963-119573	Sequence 119573, App
39	45	35.2	23	14	US-10-269-806-176	Sequence 176, App
40	45	35.2	41	14	US-10-269-806-170	Sequence 170, App
41	45	35.2	46	14	US-10-269-806-182	Sequence 182, App
42	45	35.2	46	14	US-10-269-806-188	Sequence 188, App
43	45	35.2	69	16	US-10-437-963-128768	Sequence 128768, App
44	45	35.2	79	17	US-10-425-115-333561	Sequence 333561, App
45	45	35.2	104	17	US-10-425-115-239179	Sequence 239179, App

#### ALIGNMENTS

##### RESULT 1

US-09-835-684-5  
; Sequence 5, Application US/09835684  
; Patent No. US20020019337A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/71  
; CURRENT APPLICATION NUMBER: US/09/835,684  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 447  
; TYPE: PRT  
; ORGANISM: Erwinia amylovora  
US-09-835-684-5

Query Match 100.0%; Score 128; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 3e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDQGGGGQIGDNPFLKAMKLIA 25

Db 116 ITPDQGGGGQIGDNPFLKAMKLIA 140

##### RESULT 2

US-09-880-371-5  
; Sequence 5, Application US/09880371

4



TYPE: PRT  
ORGANISM: Erwinia amylovora  
US-10-847-142-5

Query Match 100.0%; Score 128; DB 17; Length 447;  
Best Local Similarity 100.0%; Pred. No. 3e-10; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 1 ITPDGGGGQIGDNPPLKAMKLIA 25  
|||||  
Db 116 ITPDGGGGQIGDNPPLKAMKLIA 140  
|||||

RESULT 7  
US-10-109-670-10  
Sequence 10, Application US/10109670  
Publication No. US20030105283A1  
GENERAL INFORMATION:  
APPLICANT: TAKESAKO, KAZUTOH et al.  
TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN MALASSEZIA  
FILE REFERENCE: 1422-0523P  
CURRENT APPLICATION NUMBER: US/10/109,670  
CURRENT FILING DATE: 2002-04-01  
NUMBER OF SEQ ID NOS: 58  
SEQ ID NO 10  
LENGTH: 206  
TYPE: PRT  
ORGANISM: Malassezia furfur  
US-10-109-670-10

Query Match 44.5%; Score 57; DB 14; Length 206;  
Best Local Similarity 45.0%; Pred. No. 3.4;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAM 20  
:|||||:|||||:  
Db 86 LAPQSEGGQLNDGPLKQAI 105  
:|||||:|||||:

RESULT 8  
US-10-045-792-12  
Sequence 12, Application US/10045792  
Publication No. US20030003563A1  
GENERAL INFORMATION:  
APPLICANT: Vinkemeier, Uwe  
Moarefi, Ismail  
Darnell, Jr., James E.  
Kuriyan, John  
TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/045,792  
FILING DATE: 19-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/012,710  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-194  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 837 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-045-792-12

Query Match 42.2%; Score 54; DB 14; Length 837;  
Best Local Similarity 64.3%; Pred. No. 42;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GQGGQIGDNPPLK 18  
:|||||:|||||:  
Db 802 GEGGSLGSGQPLK 815  
:|||||:|||||:

RESULT 9  
US-09-845-335-3  
Sequence 3, Application US/09845335  
Patent No. US20020058266A1  
GENERAL INFORMATION:  
APPLICANT: CLOUGH, BARBARA  
APPLICANT: PREISER, PETER  
APPLICANT: WILSON, ROBERT  
TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE  
TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS  
TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS  
FILE REFERENCE: N68837B GCW PJC DP  
CURRENT APPLICATION NUMBER: US/09/845,335  
CURRENT FILING DATE: 2001-12-26  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/140,466  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-26  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 409  
TYPE: PRT  
ORGANISM: Anacystis nidulans  
US-09-845-335-3

Query Match 39.8%; Score 51; DB 9; Length 409;  
Best Local Similarity 57.1%; Pred. No. 53;  
Matches 12; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 6 QGG--GQIGDNPPLKAMKL 24  
||| |||||:|||||:  
Db 183 QGASGQKGDNPVWDKILKM 203  
||| |||||:|||||:

RESULT 10  
US-10-424-599-214915  
Sequence 214915, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684

Wed Jan 26 08:08:25 2005

```
; NAME/KEY: unsure
; LOCATION: (1)-(149)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_142992C.1.pep
US-10-425-115-231795

Query Match      38.3%; Score 49; DB 17; Length 149;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 PDGGGGGQIGDN 14
      |||:|||||:
Db      65 PDGEGGGQTDDD 76

RESULT 13
US-10-369-493-3880
; Sequence 3880, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3880
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3880

Query Match      38.3%; Score 49; DB 14; Length 483;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      3 PDGGGGGQIGDNPLK 18
      ||:||||: ||: ||:
Db      420 PEGGGSRDGDVLMK 435

RESULT 14
US-09-908-711-100
; Sequence 100, Application US/09908711
; Patent No. US20020045230A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA128
; CURRENT APPLICATION NUMBER: US/09/908,711
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US01/01360
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,867
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01344
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01345
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/764,888
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US01/01329
; PRIOR FILING DATE: 2001-01-17

; SEQ ID NO 214915
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36094C.1.pep
US-10-424-599-214915

Query Match      39.1%; Score 50; DB 15; Length 78;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      2 TPDGGGGGQIGDNPLKA 19
      |||:|||||: ||
Db      40 TPKGGGGGPRGKNPPNKA 57

RESULT 11
US-10-437-963-149121
; Sequence 149121, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149121
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49486C.1.pep
US-10-437-963-149121

Query Match      39.1%; Score 50; DB 16; Length 177;
Best Local Similarity 47.6%; Pred. No. 31;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      5 GQGGGQIGDNPLKAMKLIA 25
      |||:|||||: ||: ||:
Db      19 GVGGGRLGGGFSNEASRLAA 39

RESULT 12
US-10-425-115-231795
; Sequence 231795, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231795
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
```

```

Query Match      37.9%; Score 48.5; DB 9; Length 365;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 12; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy      2      TPDQGGGGQIGDNPLLKAMLKL 23
      ||| : ||| ||| : ||| :
Db      284      TPDKKGOTOI-PNPLLKNIILRV 304

```

Search completed: January 25, 2005, 20:58:04  
Job time : 61.3365 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 20:02:44 ; Search time 14.1827 Seconds  
(without alignments)  
169.602 Million cell updates/sec

Title: DESAI-879\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDGGGGQIGDNPLLKAMLKLI A 25  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	447	2	T18447
2	54	42.2	837	2	I57557
3	52	40.6	135	2	B49218
4	52	40.6	702	2	S48753
5	51	39.8	409	2	S04430
6	49.5	38.7	480	2	E75433
7	49	38.3	428	2	G71177
8	49	38.3	436	2	T46107
9	49	38.3	598	2	S66669
10	49	38.3	613	2	A56031
11	48	37.5	225	2	AF2140
12	47	36.7	435	2	T21052
13	47	36.7	1343	1	H64073
14	46.5	36.3	652	2	D82317
15	46.5	36.3	867	2	H70411
16	46	35.9	435	2	F84674
17	46	35.9	605	2	S01066
18	46	35.9	1134	2	D75014
19	45.5	35.5	161	2	T42006
20	45.5	35.5	694	2	S09794
21	45	35.2	206	2	JC2574
22	45	35.2	283	2	F86184
23	45	35.2	283	2	T19933
24	45	35.2	302	2	A84263
25	45	35.2	373	2	A56392
26	45	35.2	405	2	D97563
27	45	35.2	405	2	AD2784
28	45	35.2	509	2	D83651
29	45	35.2	532	2	A90037

30	45	35.2	576	2	T17842	hypothetical prote
31	45	35.2	664	1	VEHULA	lamin A - human
32	45	35.2	665	2	S27267	lamin A - rat
33	45	35.2	665	2	S28182	lamin A - mouse
34	45	35.2	1547	2	T13847	sno protein - frui
35	44	34.4	258	2	G72313	hypothetical prote
36	44	34.4	394	2	D69370	probable acyl-CoA
37	44	34.4	412	2	T32890	steroid/thyroid/re
38	44	34.4	501	2	A11031	probable membrane
39	44	34.4	535	2	S44827	F54F2.2 protein -
40	44	34.4	790	2	E48327	COI intron A prote
41	44	34.4	875	2	AE0401	alanine-tRNA ligas
42	44	34.4	890	2	F75289	alanyl-tRNA synthe
43	43.5	34.0	415	2	S60078	Runt domain contai
44	43.5	34.0	415	2	B55563	AML2a protein - hu
45	43.5	34.0	754	2	S41391	gelsolin - America

ALIGNMENTS

RESULT 1

T18447  
HrpW protein - Erwinia amylovora  
C;Species: Erwinia amylovora  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T18447  
R;Gaudriault, S.  
submitted to the EMBL Data Library, May 1998  
A;Reference number: Z18936  
A;Accession: T18447  
A;Status: preliminary; translated from GB/EMBL/DDBU  
A;Molecule type: DNA  
A;Residues: 1-447 <GAU>  
A;Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:e1293451; PIDN:CAA74  
A;Experimental source: strain CFBP1430; specific host Pomoideae  
C;Genetics:  
A;Note: hrpW

Query Match 100.0%; Score 128; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred.No. 1e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPLLKAMLKLI A 25  
Db 116 ITPDGGGGQIGDNPLLKAMLKLI A 140

RESULT 2

I57557  
DNA-Binding Protein and transcription factor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
C;Accession: I57557  
R;Quellie, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Clev  
Mol. Cell. Biol. 15, 3336-3343, 1995  
A;Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosph  
A;Reference number: I57557; MUID:95280934; PMID:7760829  
A;Accession: I57557  
A;Status: preliminary; translated from GB/EMBL/DDBU  
A;Molecule type: mRNA  
A;Residues: 1-837 <RES>  
A;Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g1008  
C;Genetics:  
A;Gene: STAT6  
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 42.2%; Score 54; DB 2; Length 837;  
Best Local Similarity 64.3%; Pred.No. 5;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 GQGGGGQIGDNPLLK 18  
:||||:|

Db 802 GEGGSLGSQLK 815

RESULT 3  
B49218  
hemagglutinin homolog pMGAL.3 - Mycoplasma gallisepticum (fragment)  
C:Species: Mycoplasma gallisepticum  
C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: B49218  
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.  
Infect. Immun. 61, 903-909, 1993  
A:Title: Molecular cloning of a member of the gene family that encodes pMGA, a hemagglutinin  
A:Reference number: A49218; MUID:93162830; PMID:8432610  
A:Accession: B49218  
A>Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-135 <MAR>  
A:Cross-references: UNIPROT:Q53304; GB:8552116; NID:g265625; PIDN:AA25398.1; PID:g265627  
A:Experimental source: S6  
A>Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBIP:125184)  
C:Genetics:  
A:Genetic code: SGC3

Query Match 40.6%; Score 52; DB 2; Length 135;  
Best Local Similarity 64.3%; Pred. No. 1.3;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGQGGQIGDNP 15  
| | | | |  
Db 57 TNPNGGGGMDNP 70

RESULT 4  
S48753  
major surface protein (clone pMGAL.3) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S48753  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;  
FEBS Lett. 352, 347-352, 1994  
A:Title: The organisation of the multigene family which encodes the major cell surface protein  
A:Reference number: S48751; MUID:95010739; PMID:7925999  
A:Accession: S48753  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-702 <MAR>  
A:Cross-references: UNIPROT:Q49498; EMBL:128424; NID:g535687; PIDN:AA62417.1; PID:g5356  
C:Genetics:  
A:Genetic code: SGC3  
A:Start codon: GTG

Query Match 40.6%; Score 52; DB 2; Length 702;  
Best Local Similarity 64.3%; Pred. No. 8.3;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGQGGQIGDNP 15  
| | | | |  
Db 57 TNPNGGGGMDNP 70

RESULT 5  
S04430  
translation elongation factor EF-Tu.A - Synecchococcus sp.  
C:Species: Synecchococcus sp.  
C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 09-Jul-2004  
C:Accession: S04430; G60663  
R:Meng, B.Y.; Shinozaki, K.; Suglura, M.  
Mol. Gen. Genet. 216, 25-30, 1989  
A:Title: Genes for the ribosomal proteins S12 and S7 and elongation factors EF-G and EF-Tu  
A:Reference number: S04426; MUID:89281486; PMID:2499762  
A:Accession: S04430  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA

A:Residues: 1-409 <MEN>  
A:Cross-references: UNIPROT:P18668  
A>Note: the source is designated as Anacystis nidulans  
R:Ludwig, W.; Weizenegger, W.; Betzl, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moellenh  
Arch. Microbiol. 153, 241-247, 1990  
A:Title: Complete nucleotide sequences of seven subbacterial genes coding for the elonga  
A:Reference number: A60663; MUID:90240875; PMID:2110445  
A:Accession: G60663  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-324, 'D', 326-409 <LUD>  
C:Genetics:  
A:Gene: tufa  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homol  
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F13-139/Domain: translation elongation factor Tu homology <ETU>  
F19-26/Region: nucleotide-binding motif A (P-loop)  
F136-139/Region: GTP-binding NXKD motif  
F174-176/Region: GTP-binding SAK/L motif  
F125,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #st

Query Match 39.8%; Score 51; DB 2; Length 409;  
Best Local Similarity 57.1%; Pred. No. 6.5;  
Matches 12; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 6 QGG--GQIGDNPFLKMLKLI 24  
| | | | | : | | |  
Db 183 QGGASGQKGDNPWDKILKLM 203

RESULT 6  
E75433  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: E75433  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; M  
S.; Smith, H.O.; Venster, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: E75433  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-480 <WHI>  
A:Cross-references: UNIPROT:Q9RV87; GB:AE001963; NID:g6458869; PIDN:AAF107  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR142  
A:Map position: 1  
C:Superfamily: Deinococcus radiodurans hypothetical protein DR142

Query Match 38.7%; Score 49.5; DB 2; Length 480;  
Best Local Similarity 60.0%; Pred. No. 13;  
Matches 12; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 5 GGGGGQIGD-NPLLKMLKLI 23  
| | | | | : | | |  
Db 157 GGGGALGDFAPLLAFRL 176

RESULT 7  
G71177  
hypothetical protein PH1701 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: G71177  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137; PMID:9679194



Search completed: January 25, 2005, 20:37:00  
Job time : 16.1827 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 19:58:54 ; Search time 79.0865 Seconds  
(without alignments)  
181.881 Million cell updates/sec

Title: DESAI-879\_116\_140  
Perfect score: 128  
Sequence: 1 ITPDGGGGQIGDNPLKXMKLKLIA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 02.4\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	447	2 O54508	O54508 erwinia amy
2	128	100.0	447	2 Q9LAW2	Q9LAW2 erwinia amy
3	128	100.0	450	2 Q6XDB8	Q6XDB8 erwinia pyr
4	128	100.0	450	2 AAQ17046	AAQ17046 erwinia p
5	128	100.0	450	2 AAS45453	AAS45453 erwinia p
6	120	93.8	138	2 Q79AW7	Q79AW7 erwinia amy
7	71	55.5	479	2 Q6RK52	Q6RK52 pectobacter
8	71	55.5	479	2 Q6D5C8	Q6D5C8 erwinia car
9	71	55.5	479	2 AAS20352	AAS20352 pectobact
10	60	46.9	237	2 Q873M4	Q873M4 malassezia
11	57	44.5	174	2 Q9P923	Q9P923 pneumocysti
12	56	43.8	165	2 Q9P920	Q9P920 pneumocysti
13	54	42.2	172	2 Q9P921	Q9P921 pneumocysti
14	54	42.2	220	2 Q74200	Q74200 pneumocysti
15	54	42.2	837	1 STA6 MOUSE	P52633 mus musculu
16	53	41.4	173	2 Q9P919	Q9P919 pneumocysti
17	52	40.6	133	2 Q49469	Q49469 mycoplasma
18	52	40.6	135	2 Q53304	Q53304 mycoplasma
19	52	40.6	210	1 SODM ASPFU	Q92450 aspergillus
20	52	40.6	229	2 Q6CFAL	Q6CFAL yarrowia li
21	52	40.6	584	2 Q9KH13	Q9KH13 mycoplasma
22	52	40.6	692	2 Q7NAN9	Q7NAN9 mycoplasma
23	52	40.6	702	2 Q49498	Q49498 mycoplasma
24	52	40.6	708	2 Q7NAP2	Q7NAP2 mycoplasma
25	52	40.6	710	2 Q7NBR1	Q7NBR1 mycoplasma
26	51	39.8	409	1 EFTU SYNPF	P18668 synchococc
27	51	39.8	409	1 EFTU SYNPF	P33171 synchococc
28	51	39.8	578	2 Q7S389	Q7S389 neurospora
29	51	39.8	667	2 Q7NBR3	Q7NBR3 mycoplasma
30	51	39.8	697	2 Q7NBS5	Q7NBS5 mycoplasma
31	50.5	39.5	504	2 Q6BXD4	Q6BXD4 debaryomyce

32 50.5 39.5 891 2 Q7VQG3 Q7VQG3 candidatus  
33 50 39.1 61 2 P70383 P70383 mus musculu  
34 50 39.1 177 2 Q9FTW3 Q9FTW3 oryza sativ  
35 50 39.1 889 1 SYA\_SYNXP Q7u3r9 synchococc  
36 50 39.1 892 2 Q7VEG5 Q7veg5 prochloroco  
37 50 39.1 1759 2 Q73XY5 Q73xy5 mycobacteri  
38 50 39.1 1759 2 AAS04490 AAS04490 mycobacte  
39 49.5 38.7 97 2 Q8LOW7 Q8LOW7 vibrio para  
40 49.5 38.7 480 2 Q9RV87 Q9rv87 deinococcus  
41 49.5 38.7 962 2 Q8DIY4 Q8diy4 synchococc  
42 49 38.3 428 2 O59360 O59360 pyrococcus  
43 49 38.3 436 2 Q9FT45 Q9ft45 arabidopsis  
44 49 38.3 452 2 Q7RV58 Q7rv58 neurospora  
45 49 38.3 593 2 Q28248 Q28248 canis famil

#### ALIGNMENTS

RESULT 1  
O54508 PRELIMINARY; PRT; 447 AA.  
AC O54508;  
DT 01-JUN-1998 (T-EMBLrel. 06, Created)  
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
DE HrpW protein.  
GN Names=hrpW;  
OS Erwinia amylovora.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Erwinia.  
OX NCBI\_TaxID=552;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CPBP1430;  
RX MEDLINE=98086111; PubMed=9426142;  
RA Gaudriault S., Malandrin L., Paulin J.P., Barny M.A.;  
RT "DepA, an essential pathogenicity factor of Erwinia amylovora showing  
RT homology with AvrE of Pseudomonas syringae, is secreted via the Hrp  
RT secretion pathway in a DapB-dependent way.";  
RL Mol. Microbiol. 26:1057-1069(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CFBP1430;  
RX MEDLINE=98316710; PubMed=9654138;  
RA Gaudriault S., Brisset M.N., Barny M.A.;  
RT "HrpW of Erwinia amylovora, a new Hrp-secreted protein.";  
RL FEBS Lett. 428:224-228(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EA321;  
RA Kim J.F., Zumoff C.H., Beer S.V.;  
RT "HrpW, a new harpin of Erwinia amylovora, is a member of a family of  
RT pectate lyases.";  
RL Phycopathology 87:0-0(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EA321;  
RA Kim J.F., Beer S.V.;  
RX MEDLINE=98422475; PubMed=9748455;  
RT "HrpW of Erwinia amylovora, a new harpin that contains a domain  
RT homologous to pectate lyases of a distinct class.";  
RL J. Bacteriol. 180:5203-5210(1998).  
DR EMBL; Y13831; CAA74158.1; -;  
DR EMBL; U94513; AAC62314.1; -;  
DR PIR; T18447; T18447.  
DR HSP; Q9RHW0; IEE6.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0005070; F:pectate lyase activity; IEA.  
DR InterPro; IPR011050; Pectin\_lyas\_like.  
DR InterPro; IPR004898; Pect\_lyase.  
DR Pfam; PF03211; Pectate\_lyase; 1.  
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

DR	InterPro; IPR011050; Pectin_lyas_like.				
DR	InterPro; IPR004898; Pect_lyase.				
DR	Pfam; PF03211; Pectate lyase; 1.				
SQ	SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;				
Query Match	100.0%;	Score 128;	DB 2;	Length 450;	
Best Local Similarity	100.0%;	Pred. No. 5.5e-10;			
Matches	25;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	114	ITPDGQGGGQIGDNPPLKAMLKLI	138		
RESULT 4					
AAQ17046					
ID	AAQ17046	PRELIMINARY;	PRT;	450 AA.	
AC	AAQ17046;				
DT	03-MAR-2004 (TrEMBLrel. 27, Created)				
DT	03-MAR-2004 (TrEMBLrel. 27, Last sequence update)				
DT	03-MAR-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	HrpW.				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,				
RA	Hur J.H., Lim C.K.;				
RT	"Molecular characterization hrp genes cluster of Erwinia pyrifoliae				
RT	and expression of hrpEp encoding elicitor of the hypersensitive				
RT	response."				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY237642; AAQ17046.1; -.				
SQ	SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;				
Query Match	100.0%;	Score 128;	DB 2;	Length 450;	
Best Local Similarity	100.0%;	Pred. No. 5.5e-10;			
Matches	25;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	114	ITPDGQGGGQIGDNPPLKAMLKLI	138		
RESULT 5					
AAS45453					
ID	AAS45453	PRELIMINARY;	PRT;	450 AA.	
AC	AAS45453;				
DT	02-MAR-2004 (TrEMBLrel. 27, Created)				
DT	02-MAR-2004 (TrEMBLrel. 27, Last sequence update)				
DT	02-MAR-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	HrpW.				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,				
RA	Hur J.H., Lim C.K.;				
RT	"Identification of hrp genes cluster and characterization of HR				
RT	elictor hrpNEP gene in Erwinia pyrifoliae."				
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY530755; AAS45453.1; -.				
SQ	SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;				
Query Match	100.0%;	Score 128;	DB 2;	Length 450;	

Query Match	100.0%;	Score 128;	DB 2;	Length 447;	
Best Local Similarity	100.0%;	Pred. No. 5.4e-10;			
Matches	25;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	116	ITPDGQGGGQIGDNPPLKAMLKLI	140		
RESULT 2					
Q9LAW2					
D	Q9LAW2	PRELIMINARY;	PRT;	447 AA.	
C	Q9LAW2;				
T	01-OCT-2000 (TrEMBLrel. 15, Created)				
T	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)				
T	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)				
E	Harpin HrpW.				
T	Name=hrpW;				
N	Erwinia amylovora.				
S	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
C	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
C	Enterobacteriaceae; Erwinia.				
C	NCBI_TaxID=552;				
N	[1]				
P	SEQUENCE FROM N.A.				
P	STRAIN=Ea246;				
C	Kim J.F., Laby R.J., Beer S.V.;				
L	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.				
L	EMBL; AF083620; AAF63402.1; -.				
R	HSP; Q9RHW0; 1EE6.				
R	GO; GO:0005576; C:extracellular; IEA.				
R	GO; GO:0030570; F:pectate lyase activity; IEA.				
R	InterPro; IPR011050; Pectin_lyas_like.				
R	InterPro; IPR004898; Pect_lyase.				
R	Pfam; PF03211; Pectate lyase; 1.				
R	SEQUENCE 447 AA; 45340 MW; 0BBAAE3871EDC2F6 CRC64;				
Q					
Query Match	100.0%;	Score 128;	DB 2;	Length 447;	
Best Local Similarity	100.0%;	Pred. No. 5.4e-10;			
Matches	25;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	116	ITPDGQGGGQIGDNPPLKAMLKLI	140		
RESULT 3					
Q6XDB8					
D	Q6XDB8	PRELIMINARY;	PRT;	450 AA.	
AC	Q6XDB8;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	Name=hrpW;				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.K.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.				
EMBL	AY237642; AAQ17046.1; -.				
DR	EMBL; AY530755; AAS45453.1; -.				
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	116	ITPDGQGGGQIGDNPPLKAMLKLI	140		
RESULT 5					
Q6XDB8					
D	Q6XDB8	PRELIMINARY;	PRT;	450 AA.	
AC	Q6XDB8;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	Name=hrpW;				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.K.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.				
EMBL	AY237642; AAQ17046.1; -.				
DR	EMBL; AY530755; AAS45453.1; -.				
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	116	ITPDGQGGGQIGDNPPLKAMLKLI	140		
RESULT 5					
Q6XDB8					
D	Q6XDB8	PRELIMINARY;	PRT;	450 AA.	
AC	Q6XDB8;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	Name=hrpW;				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.K.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.				
EMBL	AY237642; AAQ17046.1; -.				
DR	EMBL; AY530755; AAS45453.1; -.				
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	116	ITPDGQGGGQIGDNPPLKAMLKLI	140		
RESULT 5					
Q6XDB8					
D	Q6XDB8	PRELIMINARY;	PRT;	450 AA.	
AC	Q6XDB8;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	Name=hrpW;				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.K.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.				
EMBL	AY237642; AAQ17046.1; -.				
DR	EMBL; AY530755; AAS45453.1; -.				
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	116	ITPDGQGGGQIGDNPPLKAMLKLI	140		
RESULT 5					
Q6XDB8					
D	Q6XDB8	PRELIMINARY;	PRT;	450 AA.	
AC	Q6XDB8;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	Name=hrpW;				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.K.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.				
EMBL	AY237642; AAQ17046.1; -.				
DR	EMBL; AY530755; AAS45453.1; -.				
QY	1	ITPDGQGGGQIGDNPPLKAMLKLI	25		
DB	116	ITPDGQGGGQIGDNPPLKAMLKLI	140		
RESULT 5					
Q6XDB8					
D	Q6XDB8	PRELIMINARY;	PRT;	450 AA.	
AC	Q6XDB8;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)				
DE	HrpW.				
GN	Name=hrpW;				
OS	Erwinia pyrifoliae.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Erwinia.				
OX	NCBI_TaxID=79967;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.K.,				
RA	Hur J.H., Lim C.K.;				
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WT#3;				
RA	Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho				

```
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMKLI 25
Db 114 ITPDGGGGQIGDNPPLKAMKLI 138

RESULT 6
Q79AW7 PRELIMINARY; PRT; 138 AA.
AC Q79AW7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW (fragment).
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ea321;
RX MEDLINE=98115919; PubMed=9448330;
RA Bogdanove A.J., Kim J.F., Wei Z., Kolchinsky P., Charkowski A.O.,
RA Conlin A.K., Collmer A., Beer S.V.;
RT "Homology and functional similarity of an hrp-linked pathogenicity
RT locus, dePEF, of Erwinia amylovora and the avirulence locus avrE of
RT Pseudomonas syringae pathovar tomato.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330 (1998).
LR EMBL; U97504; AAC04849.1; -.
DR NON TER 138 138
FT SEQUENCE 138 AA; 13788 MW; 4FE177177C74B3C6 CRC64;

Query Match 93.8%; Score 120; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMKLI 23
Db 116 ITPDGGGGQIGDNPPLKAMKLI 138

RESULT 7
Q6RK52 PRELIMINARY; PRT; 479 AA.
AC Q6RK52;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
OS Pectobacterium atrosepticum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1039;
RA Holeva M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,
RA Birch P.R.J., Toth I.K.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY496066; AAS20352.1; -.
DR InterPro; IPR011050; Pectin_lyase-like.
DR InterPro; IPR004898; Pect. lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMKLI 24
Db 166 INPTADGGGQSGNDLLKALLELI 189

RESULT 8
Q6D5C8 PRELIMINARY; PRT; 479 AA.
AC Q6D5C8;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Type III effector protein.
GN Name=hrpW; ORFNames=ECA2112;
OS Erwinia carotovora subsp. atroseptica SCRI1043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1043;
RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX950851; CAG75014.1; -.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMKLI 24
Db 166 INPTADGGGQSGNDLLKALLELI 189

RESULT 9
AAS20352 PRELIMINARY; PRT; 479 AA.
AC AAS20352;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN HRPW.
OS Pectobacterium atrosepticum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1039;
RA Holeva M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,
RA Birch P.R.J., Toth I.K.;
RL "Rapid demonstration of a role early in disease development for the
RL type III secretion system of Erwinia carotovora subsp. atroseptica
RL SCRI1039 using a pooled transposon mutation grid.";
DR Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY496066; AAS20352.1; -.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMKLI 24
Db 166 INPTADGGGQSGNDLLKALLELI 189
```

```

RESULT 10
Q973M4 PRELIMINARY; PRT; 237 AA.
AC Q973M4; 2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese superoxide dismutase (Fragment).
GN Name=MnSOD;
OS Malassezia sympodialis.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Exobasidiomycetidae; Malasseziales; Malassezia.
OX NCBI_TaxID=76777;
RN [1]
RP SEQUENCE FROM N.A.
RA Anderson A., Rasool O., Schmidt M., Kodzius R., Cramer R.,
RA Scheynius A.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AJ548421; CAD68071.1; -.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON_TER 1
FT SEQUENCE 237 AA; 26716 MW; 6C2096462E8402E9 CRC64;
SQ
Query Match 46.9%; Score 60; DB 2; Length 237;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPFLKMLK 22
Db 118 MAPQSGGGQLNDGFLKQAIDK 139

RESULT 11
Q9P923 PRELIMINARY; PRT; 174 AA.
AC Q9P923;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese-dependent superoxide dismutase (EC 1.15.1.1)
DE (Fragment).
GN Name=soda;
OS Pneumocystis carinii f. sp. muris.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=42066;
RN [1]
RP SEQUENCE FROM N.A.
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferragut C., Viscogliosi E.,
RA Dei-Cas E., Wakefield A.E.;
RL "Genetic divergence at the SODA locus of six different formae
RT specialises of Pneumocystis carinii.";
RL Med. Mycol. 38:289-300(2000).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF146754; AAP25726.1; -.
DR HSP; Q92450; IKKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON_TER 1
FT SEQUENCE 174 AA; 20198 MW; 2730D8F435576124 CRC64;
SQ
Query Match 44.5%; Score 57; DB 2; Length 174;
Best Local Similarity 50.0%; Pred. No. 4.5;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPFLKMLK 22
Db 71 LLPNKGQGGQLIINGPLVEAIKK 92

RESULT 12
Q9P920 PRELIMINARY; PRT; 165 AA.
AC Q9P920;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese-dependent superoxide dismutase (EC 1.15.1.1)
DE (Fragment).
GN Name=soda;
OS Pneumocystis carinii f. sp. macaca.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=112250;
RN [1]
RP SEQUENCE FROM N.A.
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferragut C., Viscogliosi E.,
RA Dei-Cas E., Wakefield A.E.;
RL "Genetic divergence at the SODA locus of six different formae
RT specialises of Pneumocystis carinii.";
RL Med. Mycol. 38:289-300(2000).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF146754; AAP25726.1; -.
DR HSP; Q92450; IKKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON_TER 1
FT SEQUENCE 165 AA; 19178 MW; 4C01F3C7C48D514E CRC64;
SQ
Query Match 43.8%; Score 56; DB 2; Length 165;
Best Local Similarity 50.0%; Pred. No. 5.9;

```

```

Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMLK 22
: ||||| |||: |||
Db 62 LLPKGGGGQIFDGLVDAIKK 83

RESULT 13
Q9P921 PRELIMINARY; PRT; 172 AA.
AC Q9P921;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese-dependent superoxide dismutase (SC 1.15.1.1)
DN (Fragment).
GN Name=soda;
OS Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OC NCBI_TaxID=42068;
RN [1]
RP SEQUENCE FROM N.A.
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferragut C.,
RA Dei-Cas E., Wakefield A.E.;
RT "Genetic divergence at the SODA locus of six different formae
RL Med. Mycol. 38:289-300(2000).";
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
EMBL; AF146753; AAF25725.1; -.
DR HSSP; Q92450; 1KKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR PRINTS; PR01703; MNSODISMTASE.
DR PRODOM; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
DR NON_TER 172 172
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 19884 MW; CDEC754E7F7F8F2A CRC64;

Query Match 42.2%; Score 54; DB 2; Length 172;
Best Local Similarity 45.5%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMLK 22
: ||||| |||: |||
Db 69 LLPKGGGGQIFDGLVDAIKK 90

RESULT 14
O74200 PRELIMINARY; PRT; 220 AA.
ID O74200;
AC O74200;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese superoxide dismutase precursor.
DN Name=mnsod;
OS Pneumocystis carinii.
OS Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OC NCBI_TaxID=4754;

```

```

[1]
RN SEQUENCE FROM N.A.
RP Denis C.M., Guyot K., Dei-Cas E., Camus D., Odberg-Ferragut C.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
EMBL; AF036321; AAC24764.1; -.
DR HSSP; Q92450; 1KKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR000834; Peptidase_M14.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMTASE.
DR PRODOM; PD000475; SODismutase; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN 1.
DR PROSITE; PS00088; SOD_MN; 1.
DR Oxidoreductase; Signal.
KW SIGNAL 1 27 Potential.
FT CHAIN 48 169 manganese superoxide dismutase.
FT CHAIN 170 183 manganese superoxide dismutase.
FT CHAIN 184 207 manganese superoxide dismutase.
FT CHAIN 208 220 manganese superoxide dismutase.
FT CHAIN 28 35 manganese superoxide dismutase.
FT CHAIN 36 47 manganese superoxide dismutase.
SQ SEQUENCE 220 AA; 25869 MW; 73B1F11C98929E18 CRC64;

Query Match 42.2%; Score 54; DB 2; Length 220;
Best Local Similarity 45.5%; Pred. No. 16;
Matches 10; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ITPDGGGGQIGDNPPLKAMLK 22
: ||||| |||: |||
Db 106 LLPKGGGGQIFDGLVDAIKK 127

RESULT 15
STA6 MOUSE STANDARD; PRT; 837 AA.
ID STA6 MOUSE
AC P52633;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and transcription activator 6.
GN Name=Stat6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95280934; PubMed=7760829;
RA Queller F.W., Shimoda K., Thierfelder W., Fischer C.L., Kim A.,
RA Ruben S.M., Cleveland J.L., Pierce J.H., Keegan A.D., Nelms K.,
RA Paul W.E., Ihle J.N.;
RT "Cloning of murine Stat6 and human Stat6, Stat proteins that are
RT tyrosine phosphorylated in responses to IL-4 and IL-3 but are not
RT required for mitogenesis";
RL Mol. Cell. Biol. 15:3336-3343(1995).
CC -1- FUNCTION: Carries out a dual function: signal transduction and
CC activation of transcription. Involved in interleukin-4 signalling.
CC -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC member (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC in response to phosphorylation.
CC -1- PTM: Tyrosine phosphorylated following stimulation by IL-4 and IL-
CC 3.

```

Wed Jan 26 08:08:25 2005

```
CC -I- SIMILARITY: Belongs to the transcription factor STAT family.
CC -I- SIMILARITY: Contains 1 SH2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L47650; AAA79006.1; --
CC PIR; I57557; I57557.
CC HSP; P42227; IBG1.
CC MGD; MGI:103034; Stat6.
CC GO; GO:0005737; C:cytoplasm; IDA.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0042127; P:regulation of cell proliferation; IDA.
CC InterPro; IPR008967; P53_like_DNA_bnd.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR01217; STAT.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF01017; STAT_alpha; 1.
CC Pfam; PF02864; STAT_bind; 1.
CC Pfam; PF02865; STAT_int; 1.
CC SMART; SM00252; SH2; 1.
CC PROSITE; PS0001; SH2; 1.
CC Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW Transcription regulation.
FT DOMAIN 517 632 SH2.
FT MOD_RES 641 641 Phosphotyrosine (by JAK) (By similarity).
SQ SEQUENCE 837 AA; 93725 MW; FD7D7C448743EFC7 CRC64;

Query Match 42.2%; Score 54; DB 1; Length 837;
Best Local Similarity 64.3%; Pred.No. 64;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 GCGGGQIGDNPLK 18
DB 802 GCGGSLGSGPLK 815
|:|:|:|:|
|:|:|:|:|
```

Search completed: January 25, 2005, 20:35:55  
Job time : 80.0865 secs